Èdouard Lucas:

The theory of recurrent sequences is an inexhaustible mine which contains all the properties of numbers; by calculating the successive terms of such sequences, decomposing them into their prime factors and seeking out by experimentation the laws of appearance and reproduction of the prime numbers, one can advance in a systematic manner the study of the properties of numbers and their application to all branches of mathematics.

Preview

Multiresolution Redistricting

Daryl DeFord

MIT – CSAIL
Geometric Data Processing Group

Quantitative Investigations of Gerrymandering and Redistricting
Duke University
Durham, NC
March 2, 2020

Outline

MORAL #1:

MORAL #1:

Computational Redistricting is NOT a solved problem!

MORAL #2:

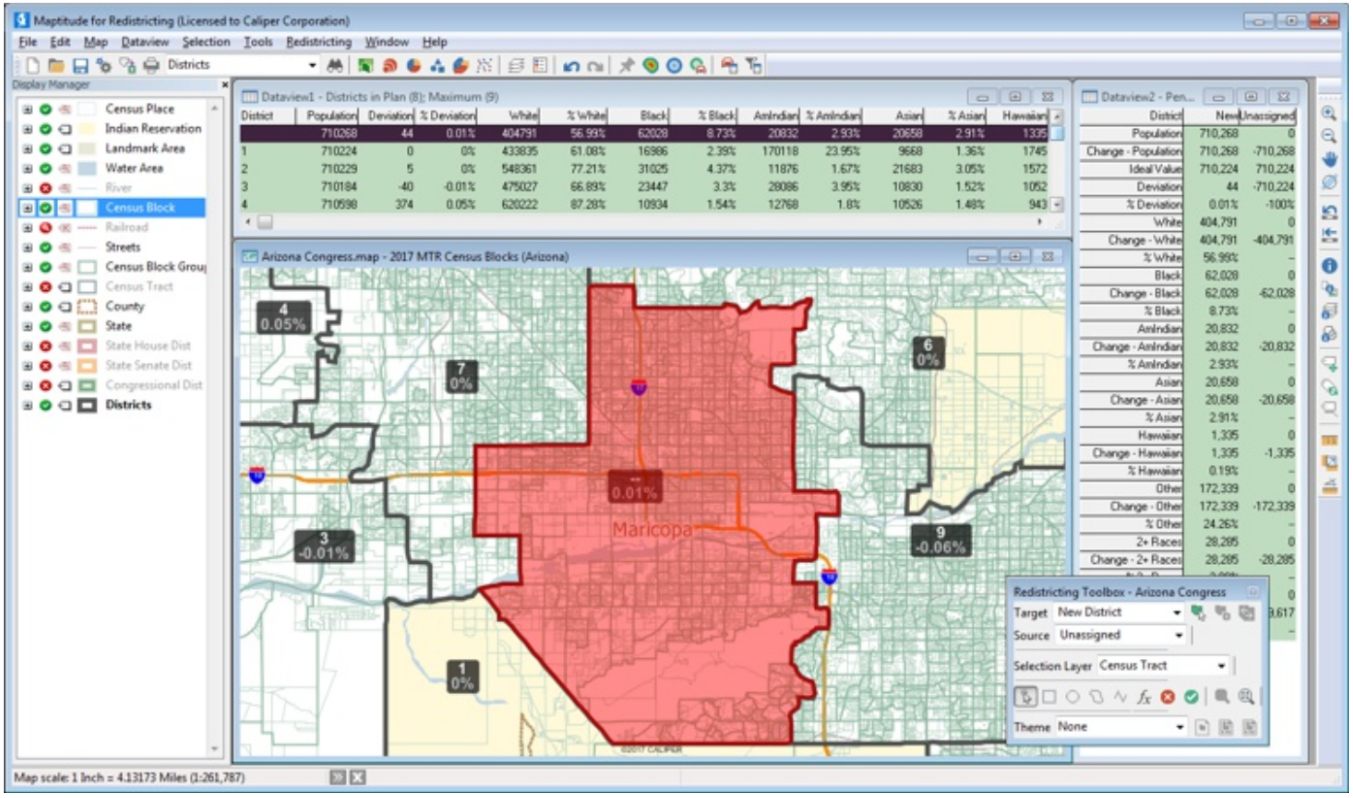
MORAL #2:

Computational Redistricting is NOT a solved problem!

MORAL #2:

Computational Redistricting is NOT a solved problem!

Ensembles in Maptitude



Potential Input Categories

- Report styles and outputs
- Methodology
- Metrics to evaluate
- Data decisions
- Communication

Target Audiences (and framework)

- Consultants
- Commissions/Legislators
- Public outreach
- Researchers

Target Audiences (and framework)

- Consultants
- Commissions/Legislators
- Public outreach
- Researchers

Current Structure:

1. Build a map
2. Push a button
3. Get a report

Possible Outputs

- Binary (p-value?)
- Distribution of summary statistics:
  - Full ensemble
  - Extreme values
  - Variants with similar properties
- Example plans:
  - Full Assignment Files
  - Filtered subsets
  - Nearby alternatives

Methodology and Implementation

- Markov chains at all?
- Target Distributions
- Sampling Techniques
- Constraints and Parameters
- Underlying Data

Methodology and Implementation

- Markov chains at all?
- Target Distributions
- Sampling Techniques
- Constraints and Parameters
- Underlying Data

Alternative Workflows:

- Local optimization/comparison
- Iterative design process
- Generate and smooth

What to measure?

- Compactness scores
- Partisan Statistics
- VRA "compliance"

What to measure?

- Compactness scores
- Partisan Statistics
- VRA "compliance"

Then what?

- Optimization
- Interactions and Tradeoffs
- Normative Concerns
- Meaningfulness

Landscape Impact

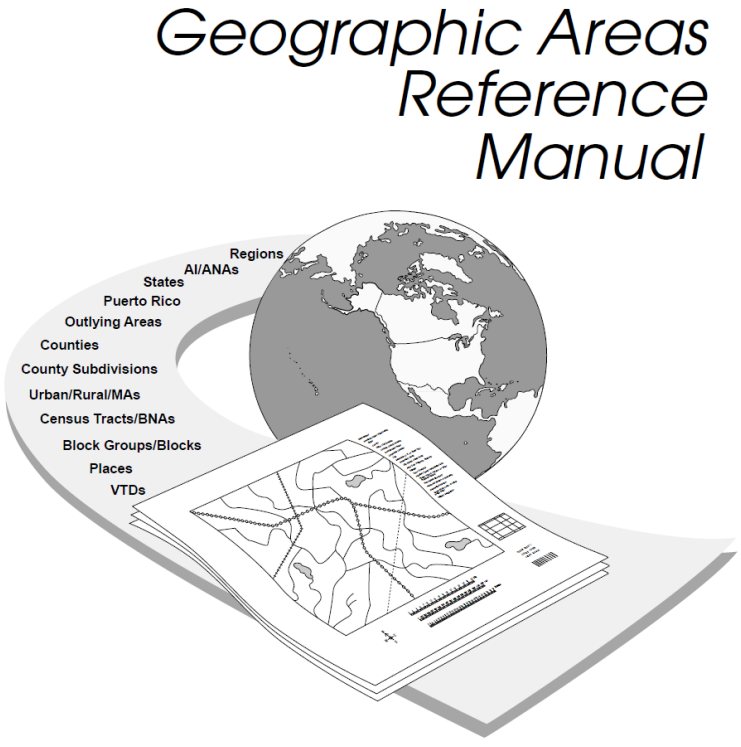
- (caveat) Proprietary Methodology

Landscape Impact

- (caveat) Proprietary Methodology
- How does this impact priorities for analysis?
- How might this change consensuses around ensembles work?
- ...
- ...
- ...

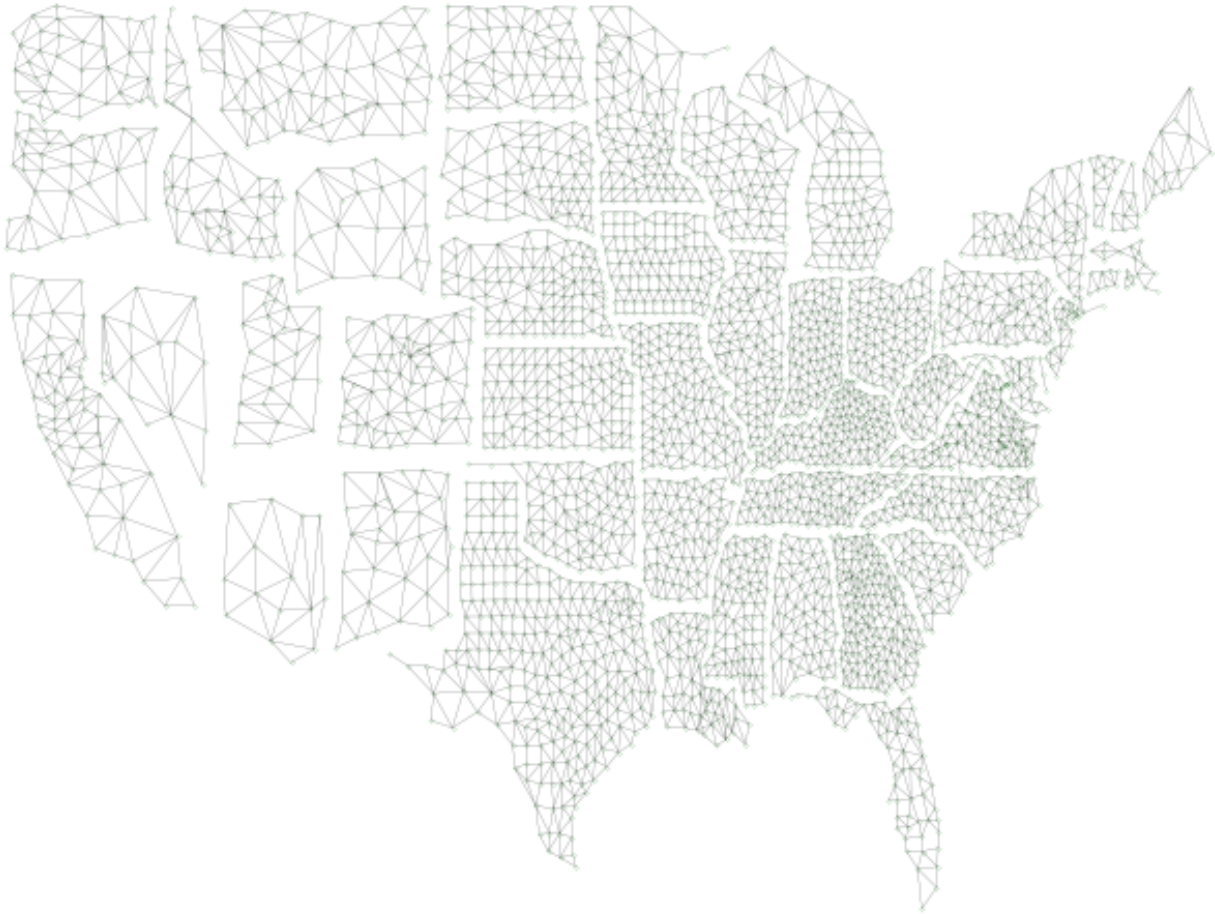
?

Census Data

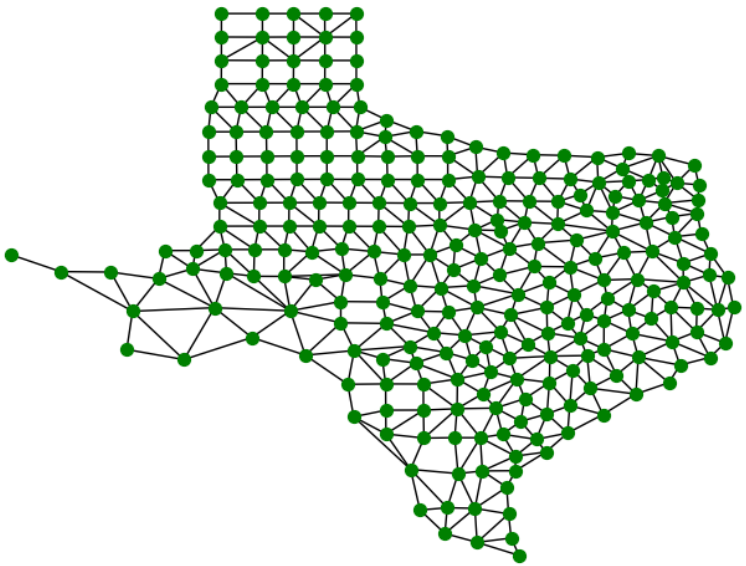


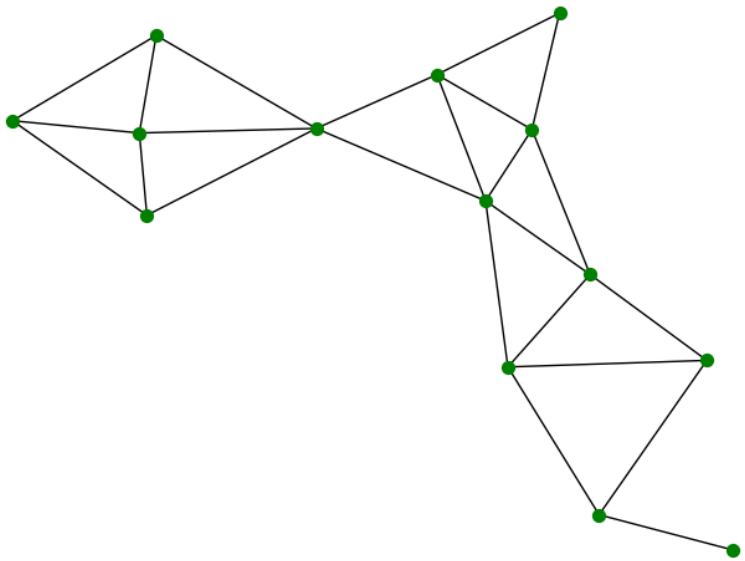
Documentation:
`https://www2.census.gov/geo/pdfs/reference/GARM/`
Data: `daryldeford.com/dual_graphs`

Counties

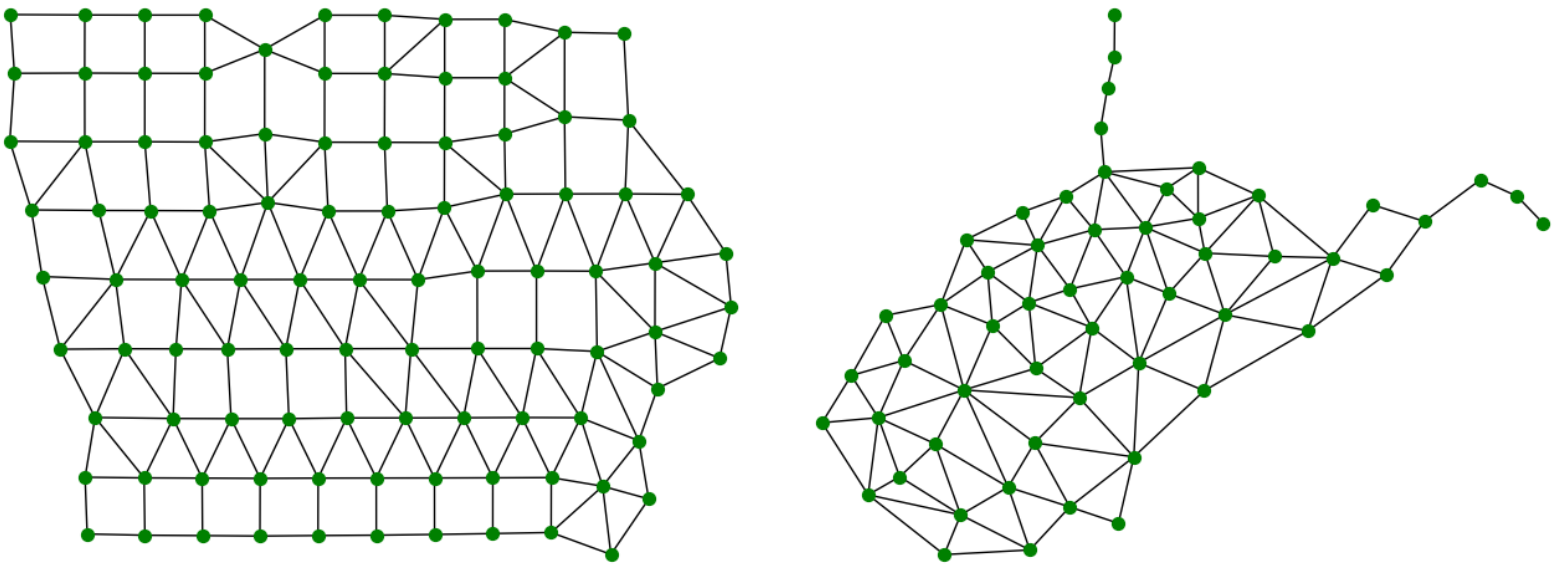


Counties

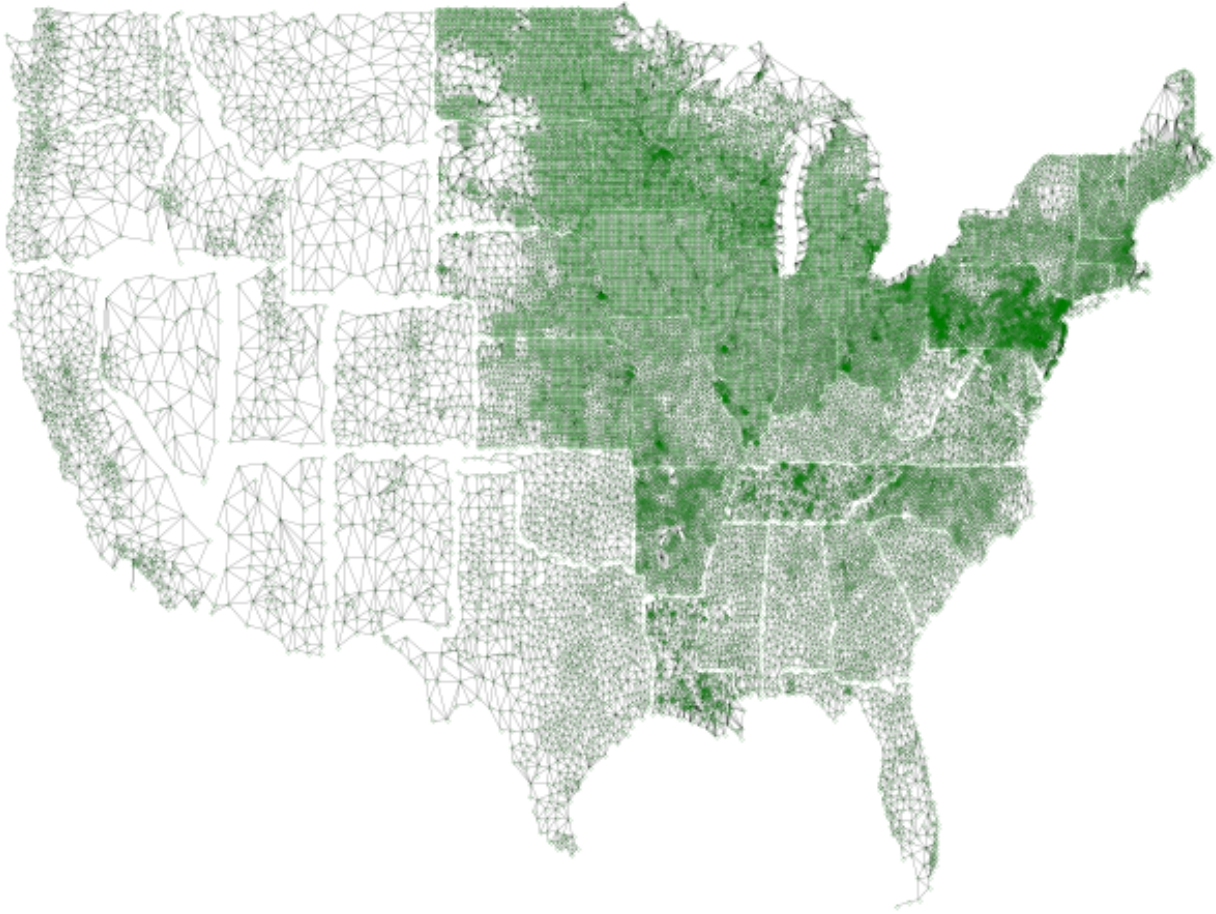




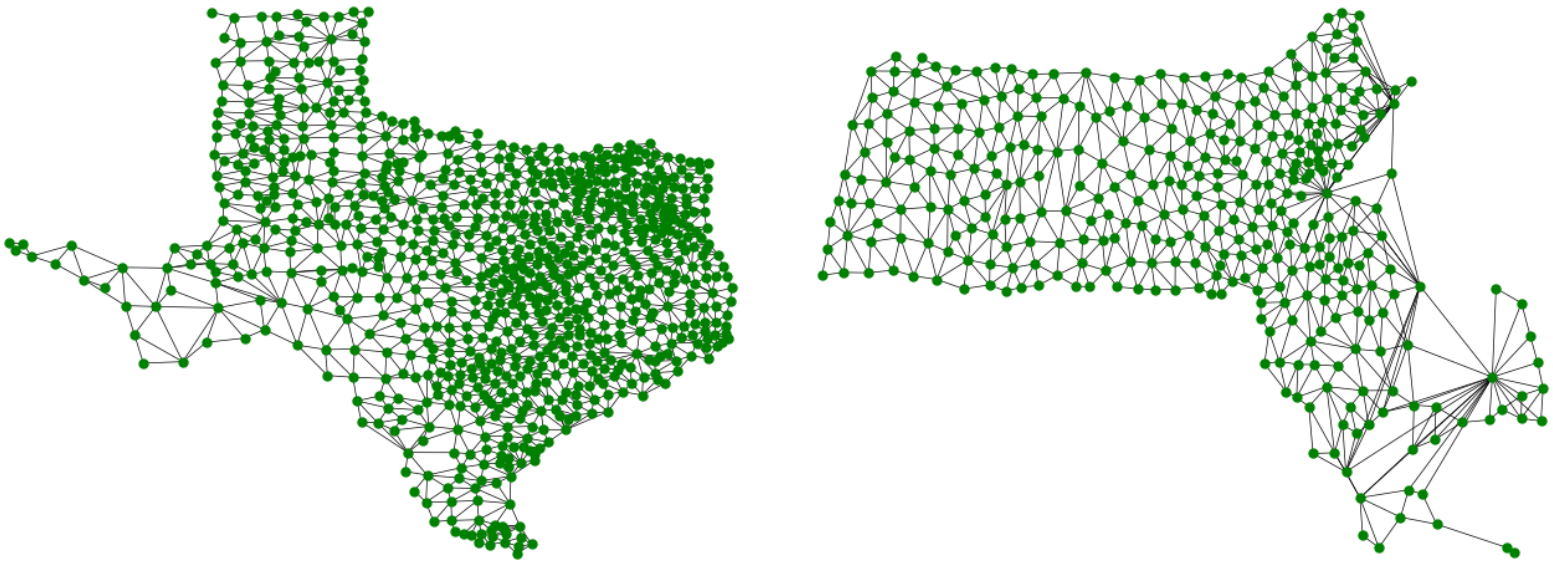
Counties



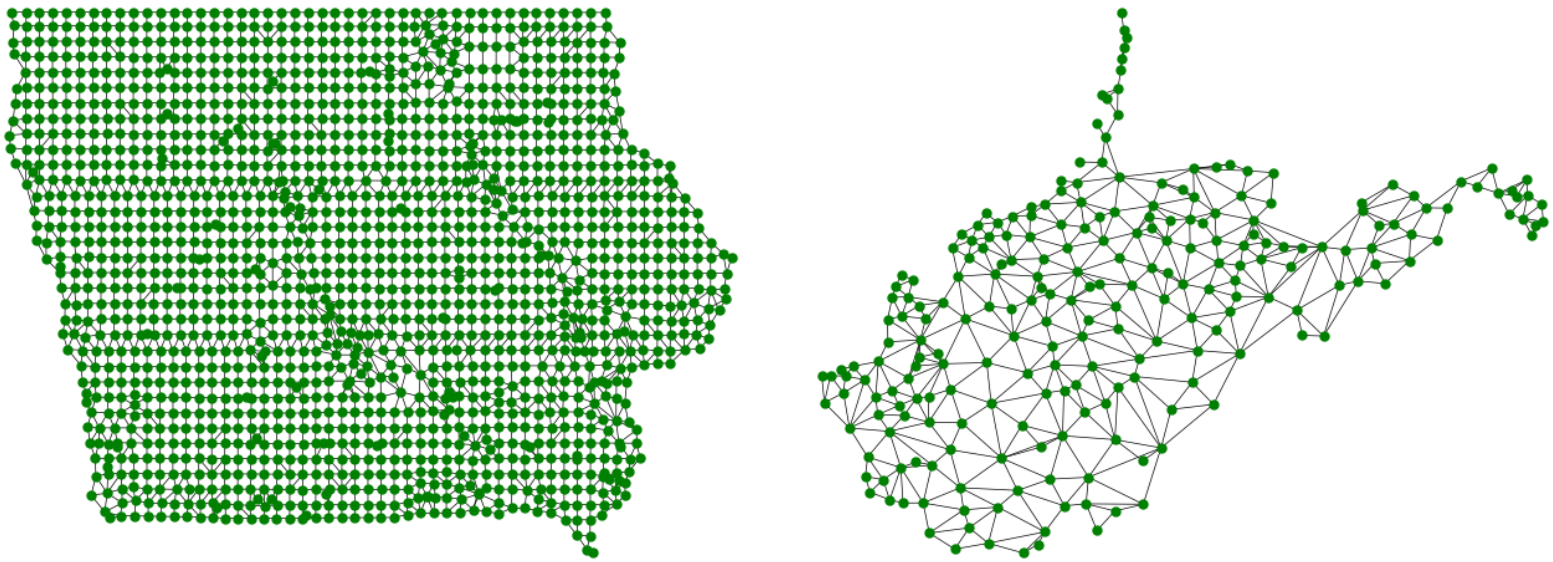
County Subunits



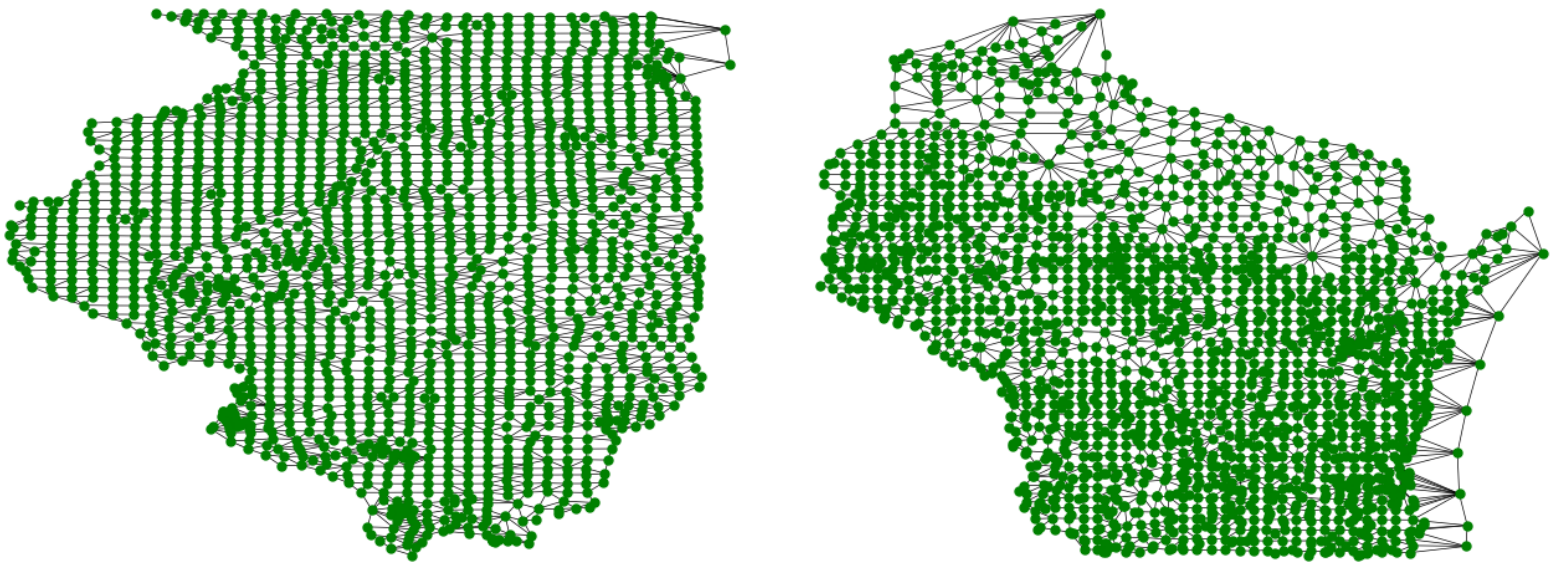
County Subunits



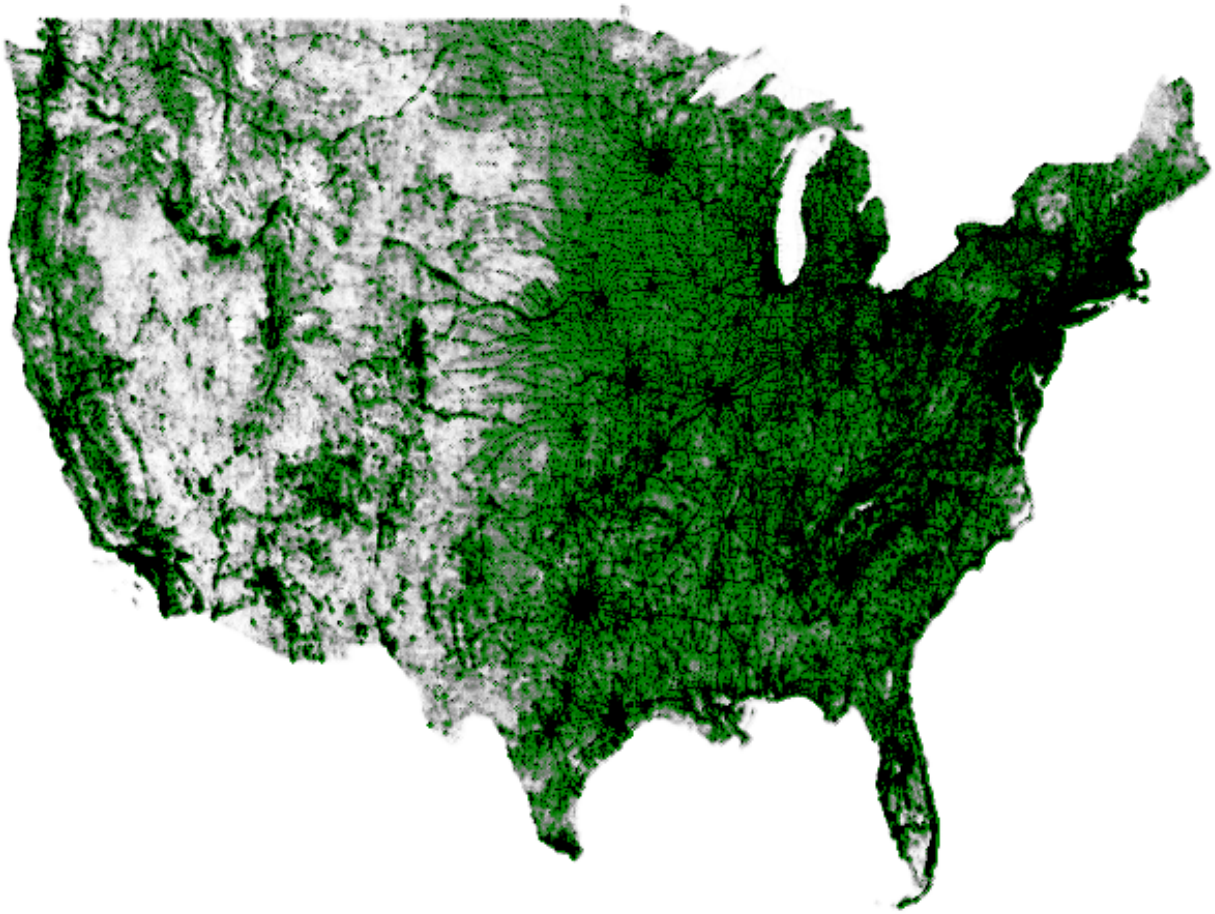
County Subunits



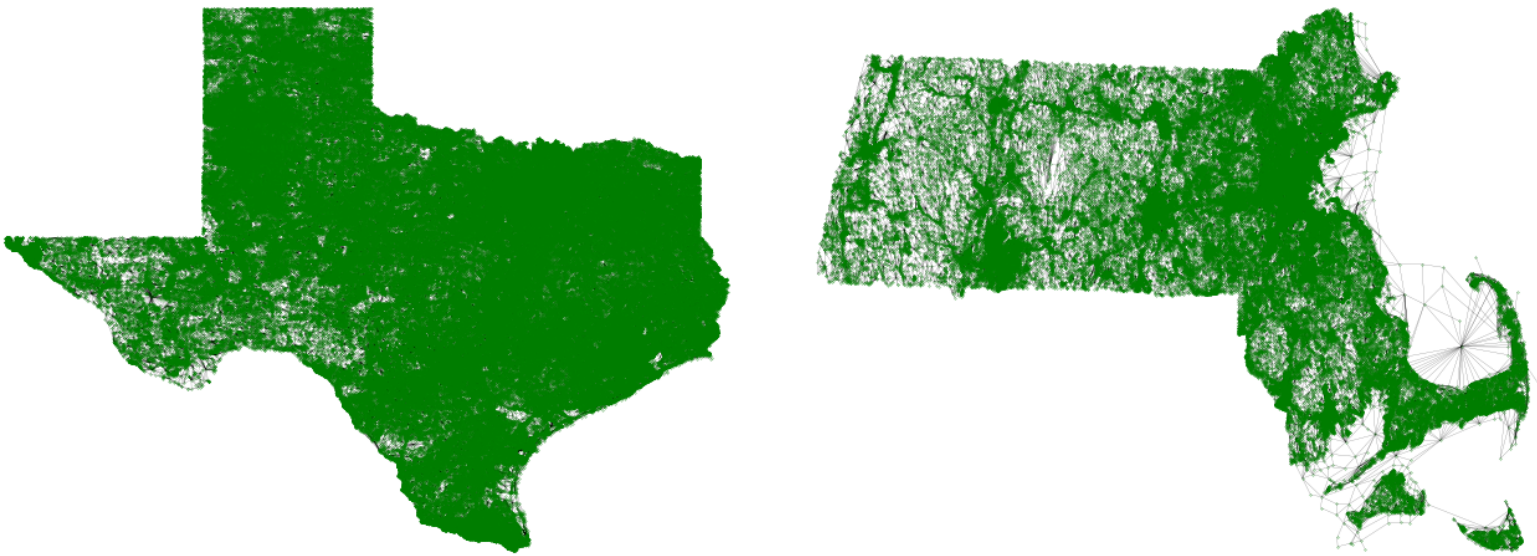
County Subunits



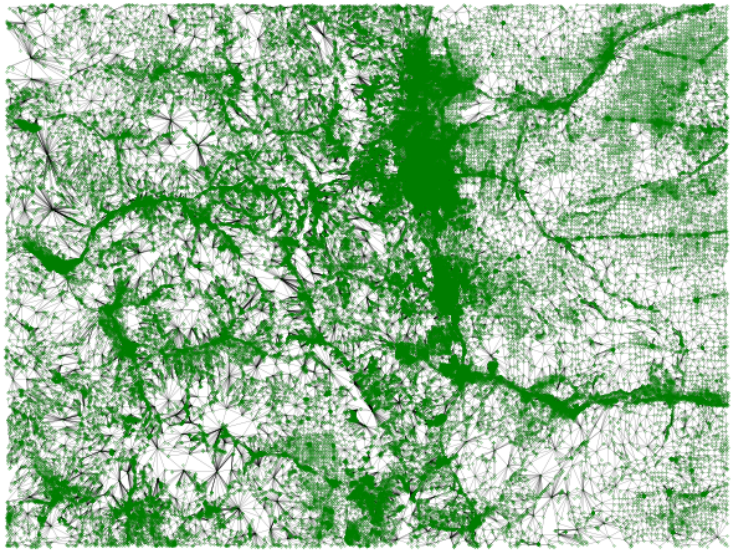
Blocks

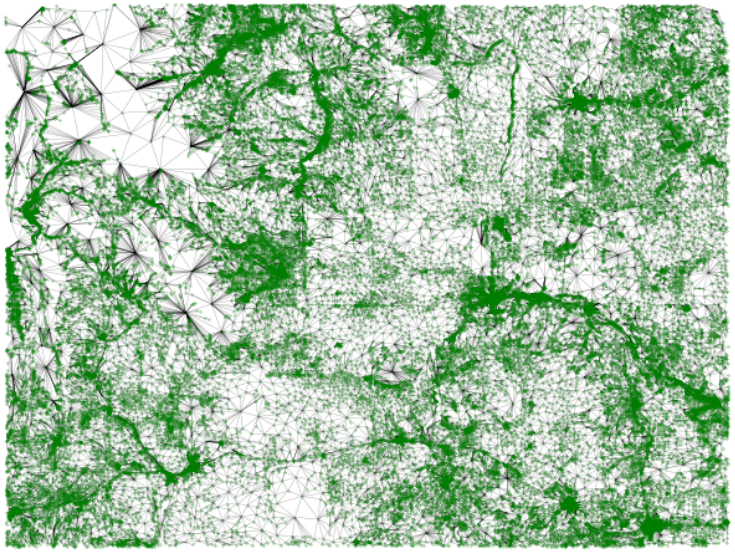


Blocks

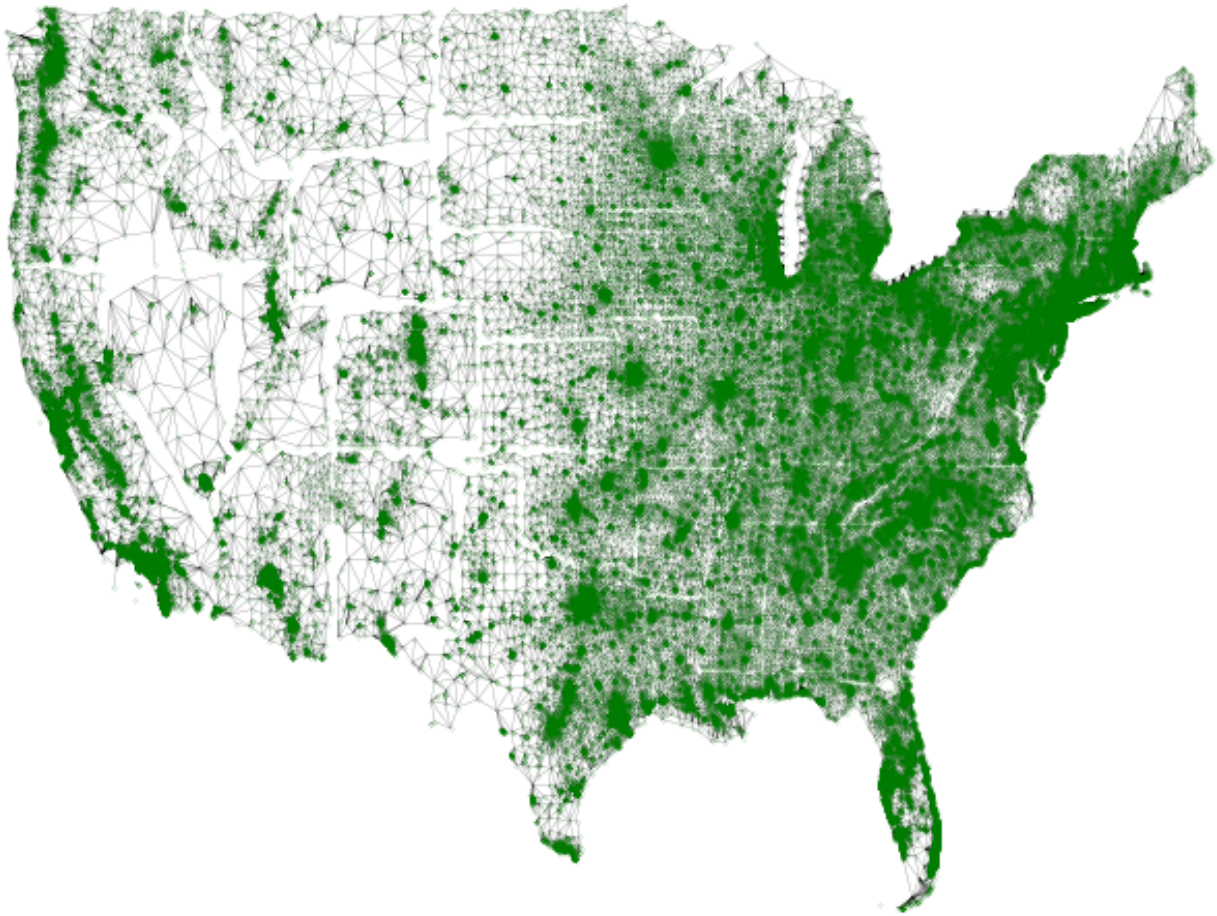


Blocks

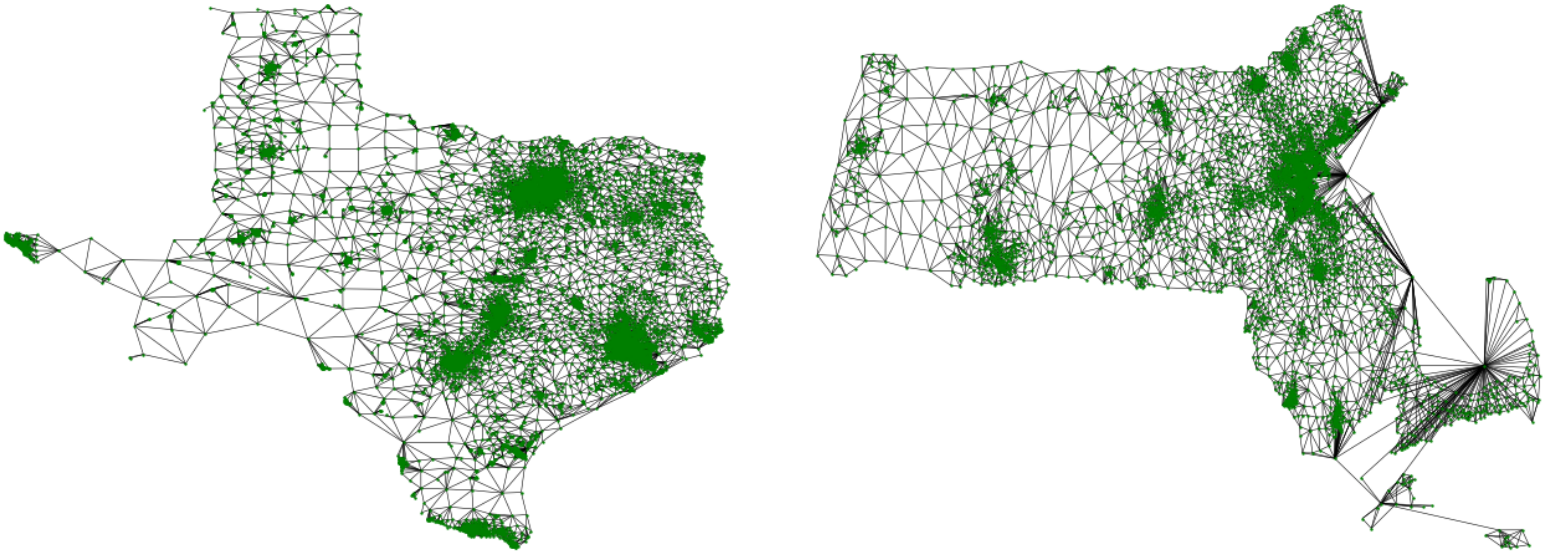




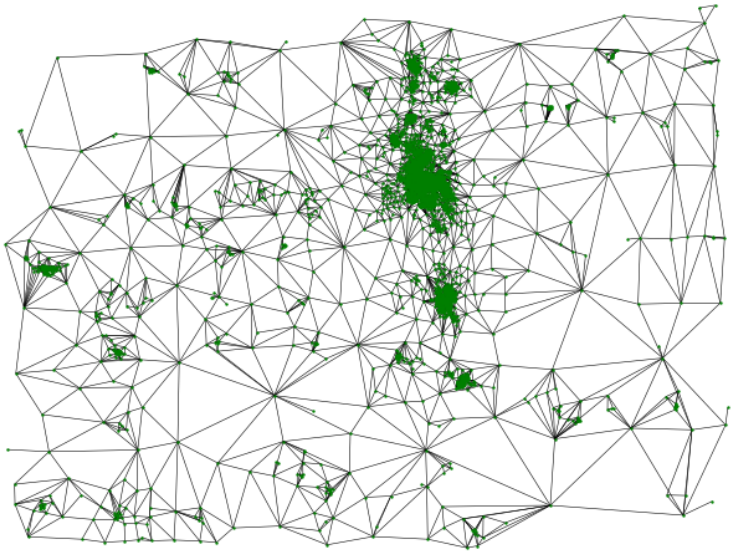
Block Groups

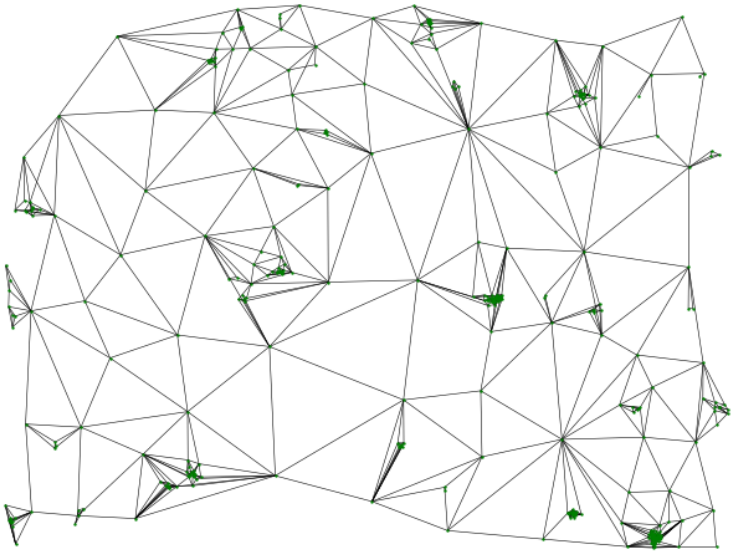


Block Groups

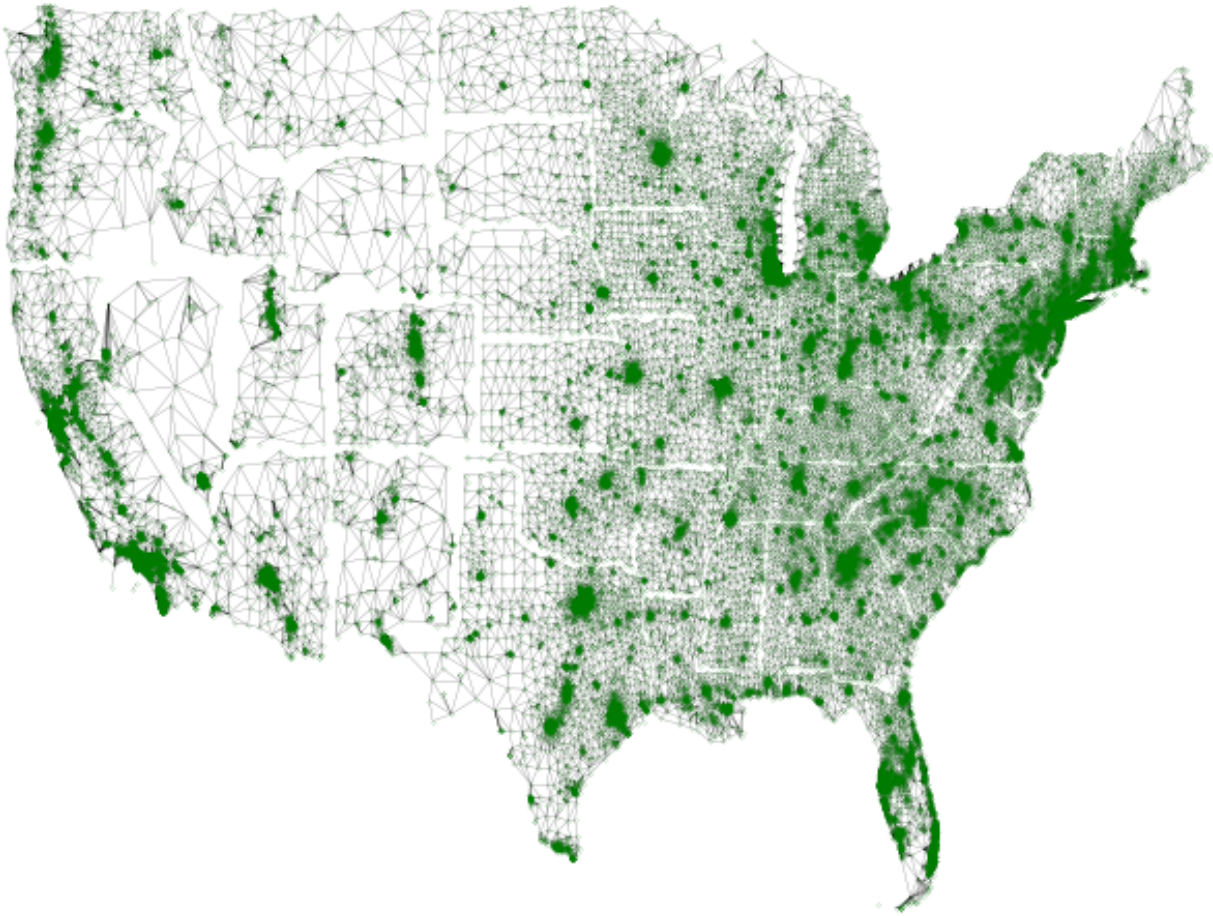


Block Groups

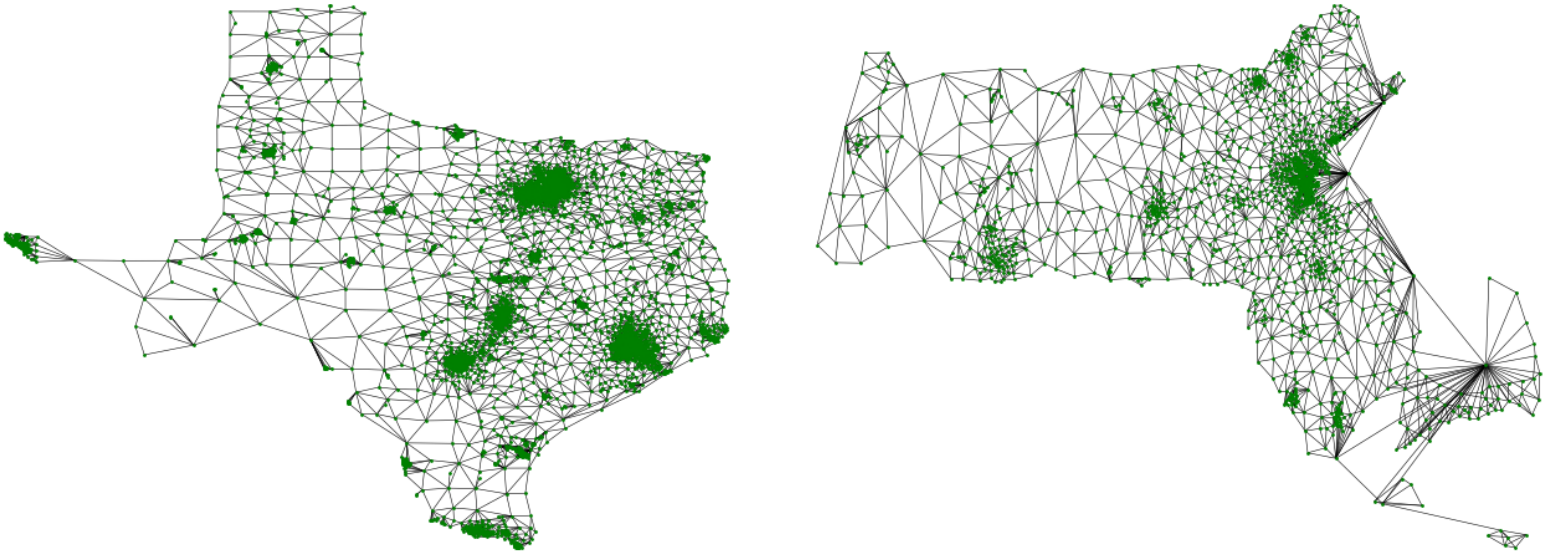




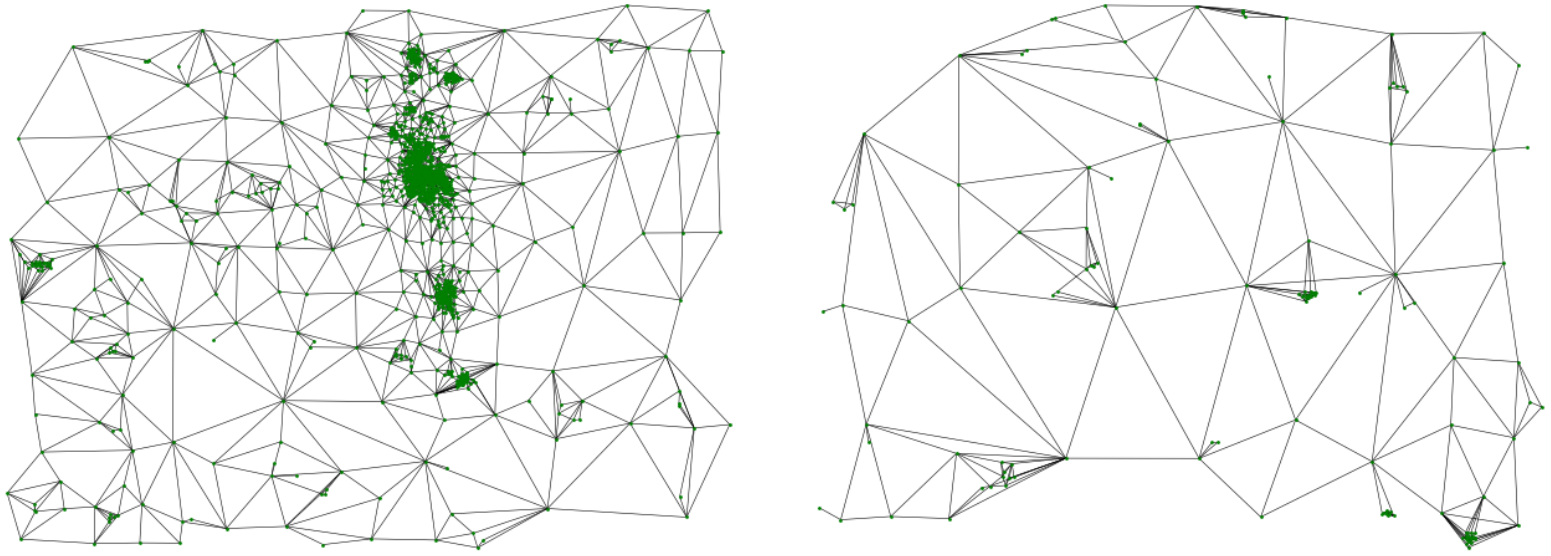
Tracts



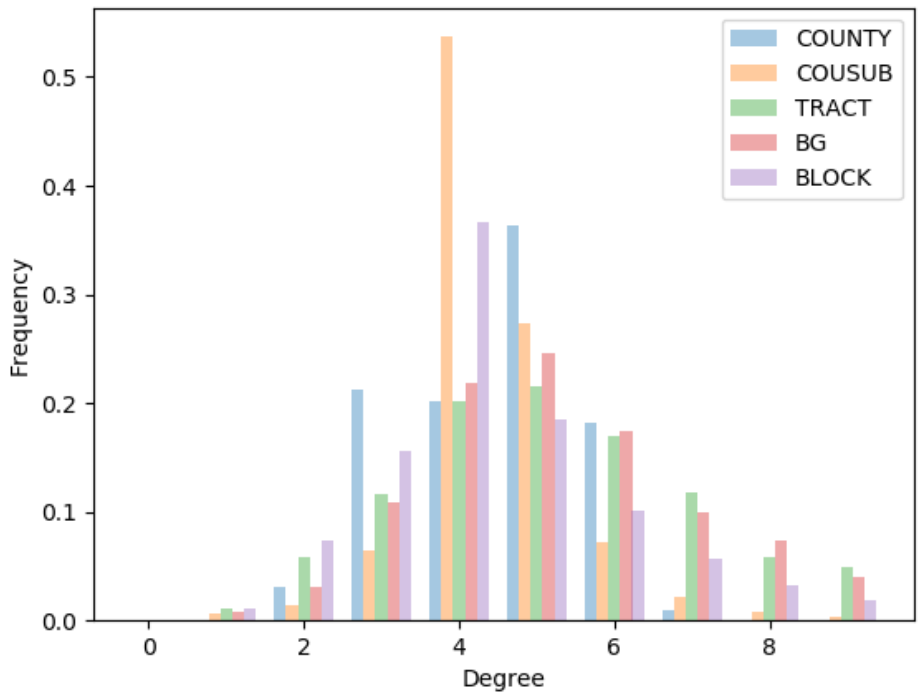
Tracts



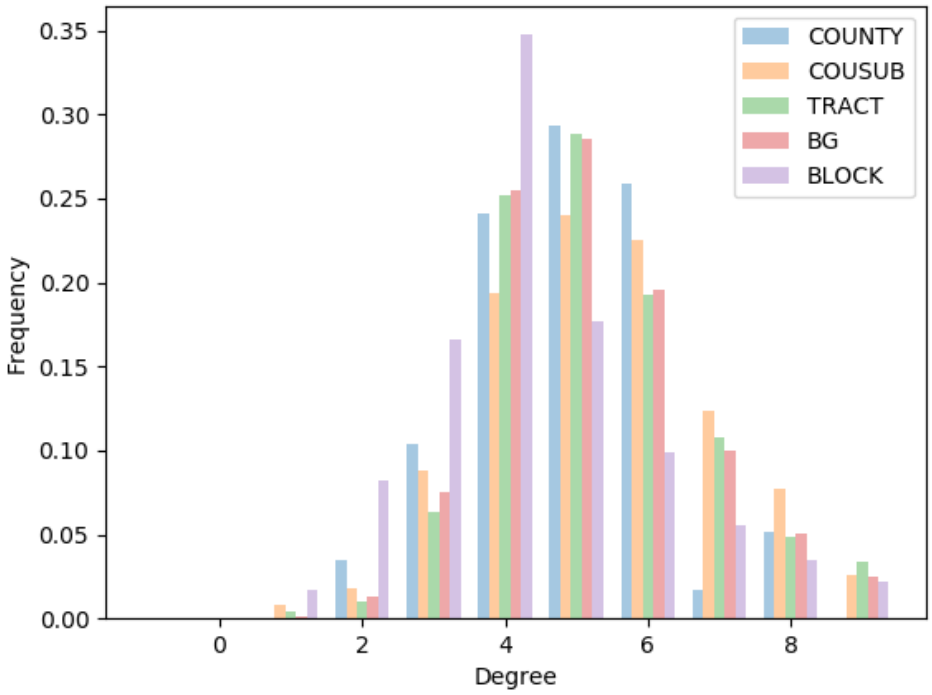
Tracts



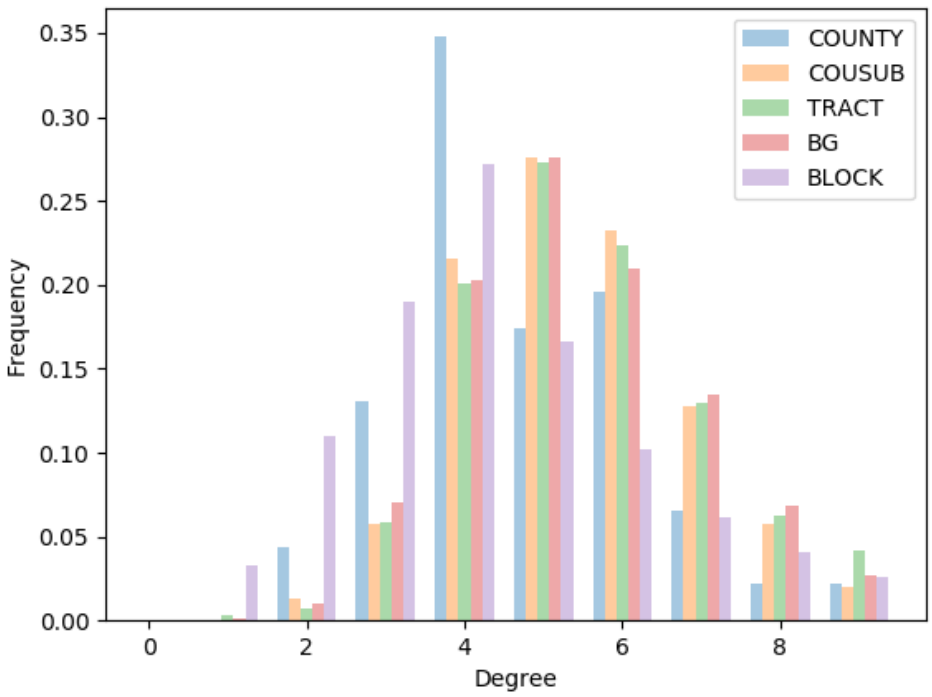
Degree Distributions (Iowa)



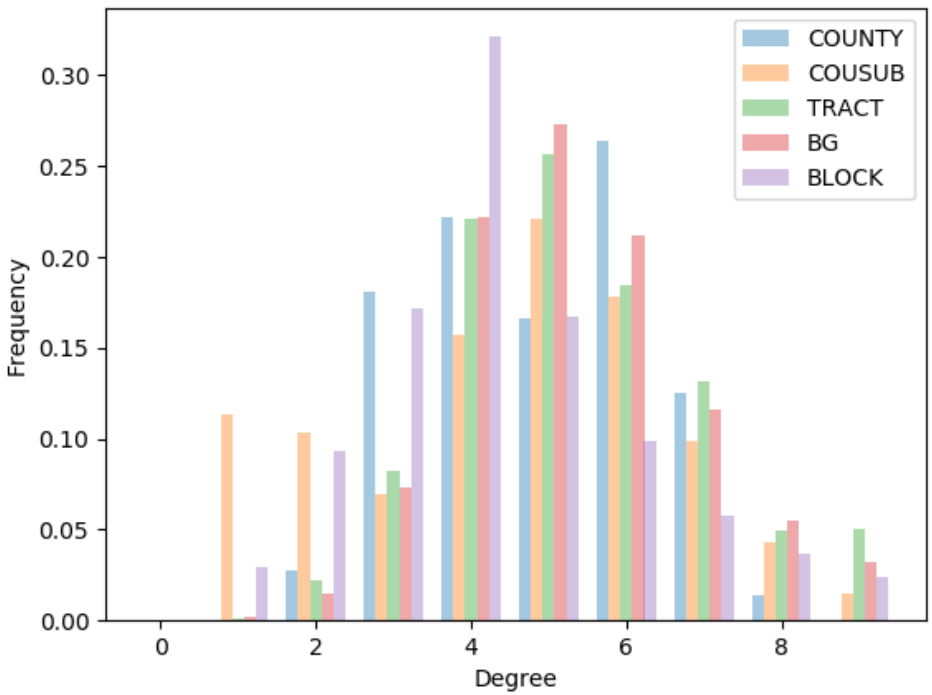
Degree Distributions (California)



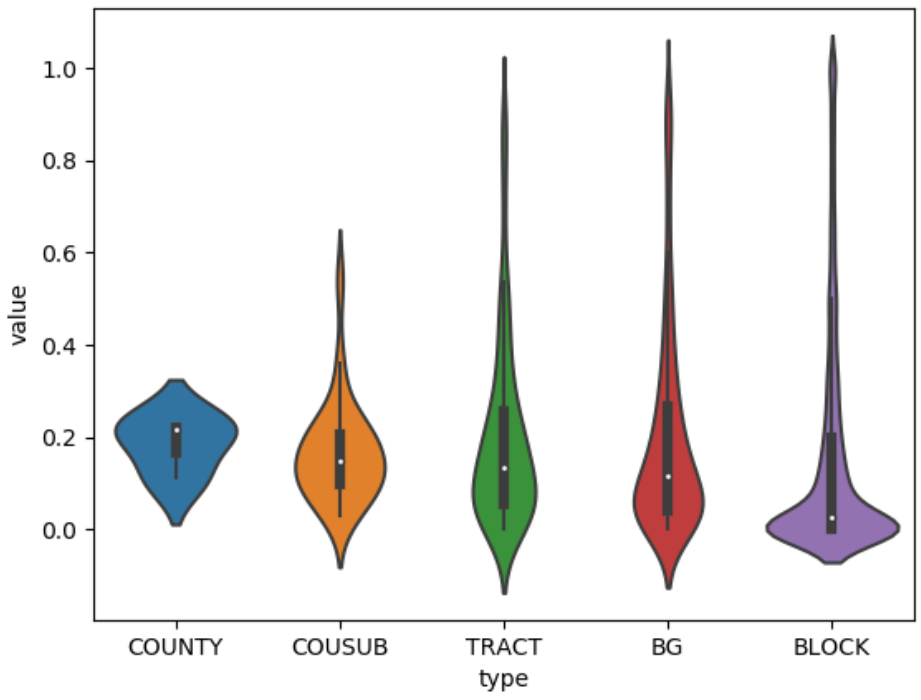
Degree Distributions (South Carolina)



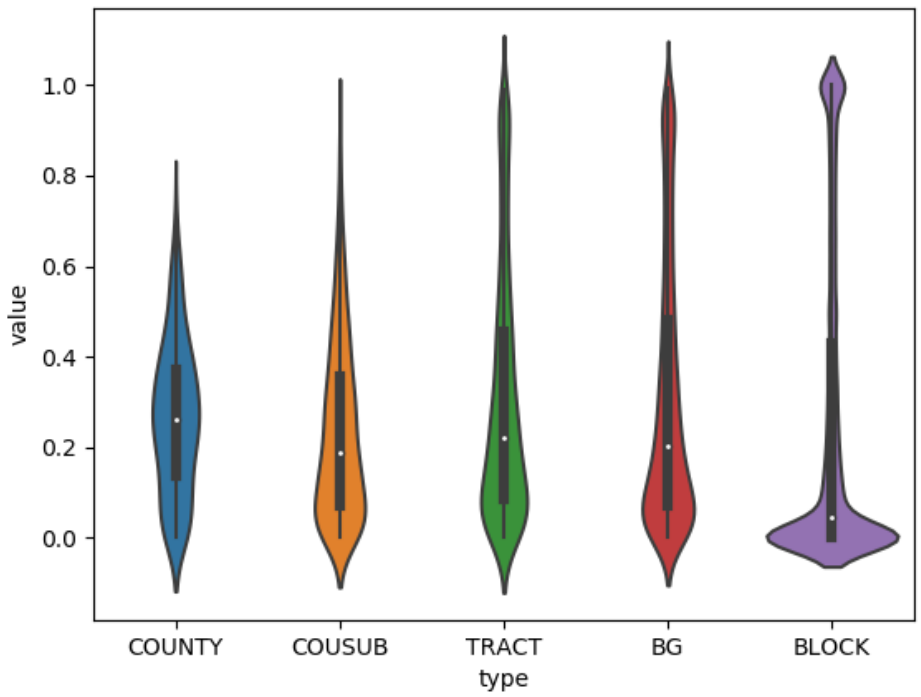
Degree Distributions (Wisconsin)



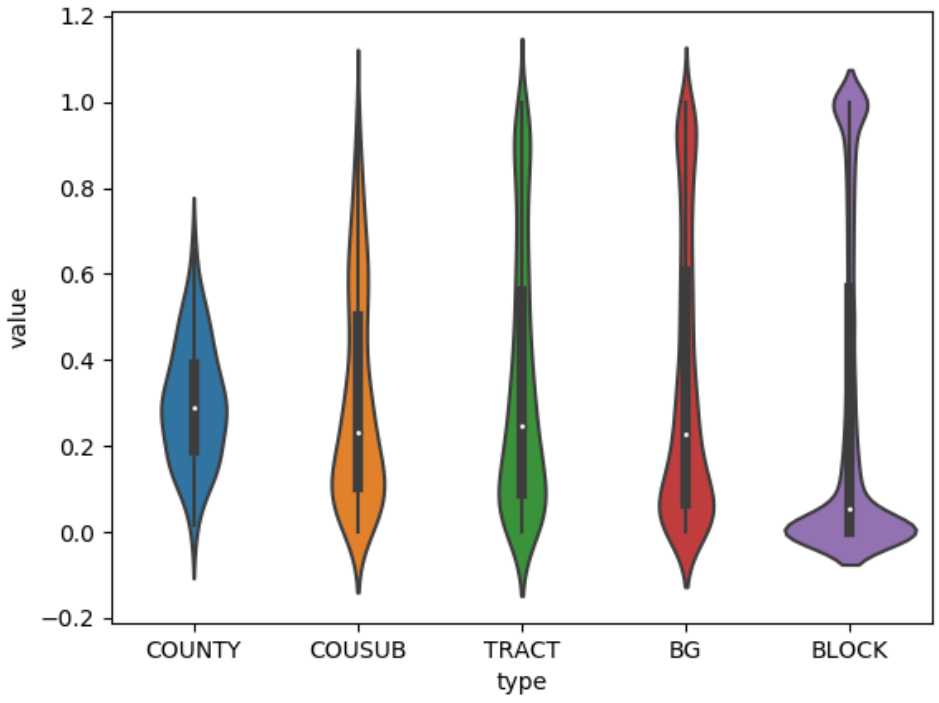
Homogeneity - BVAP - Delaware



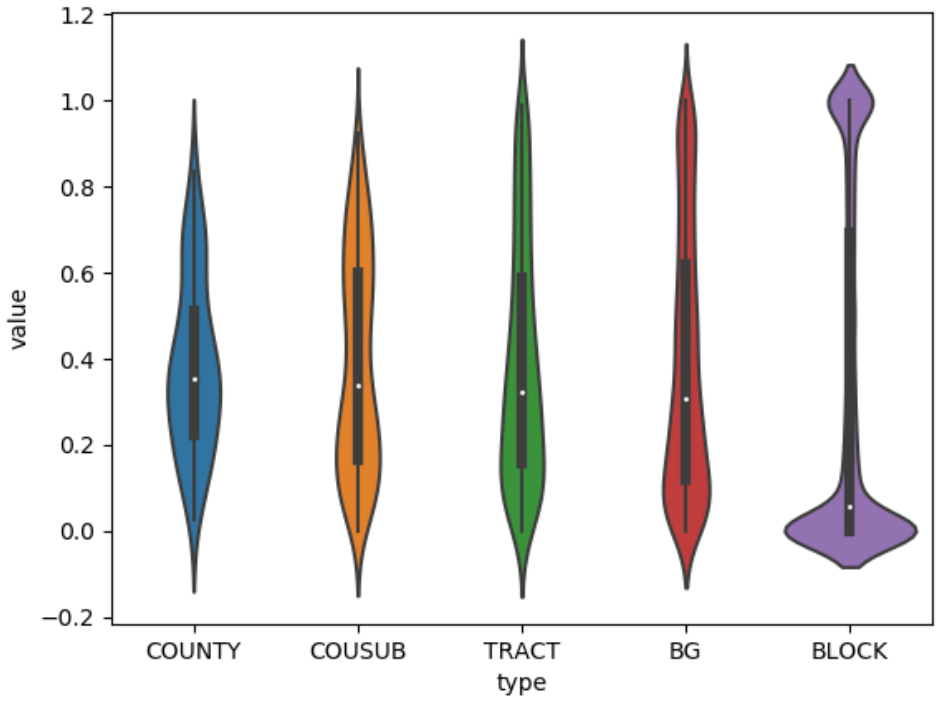
Homogeneity - BVAP - Georgia



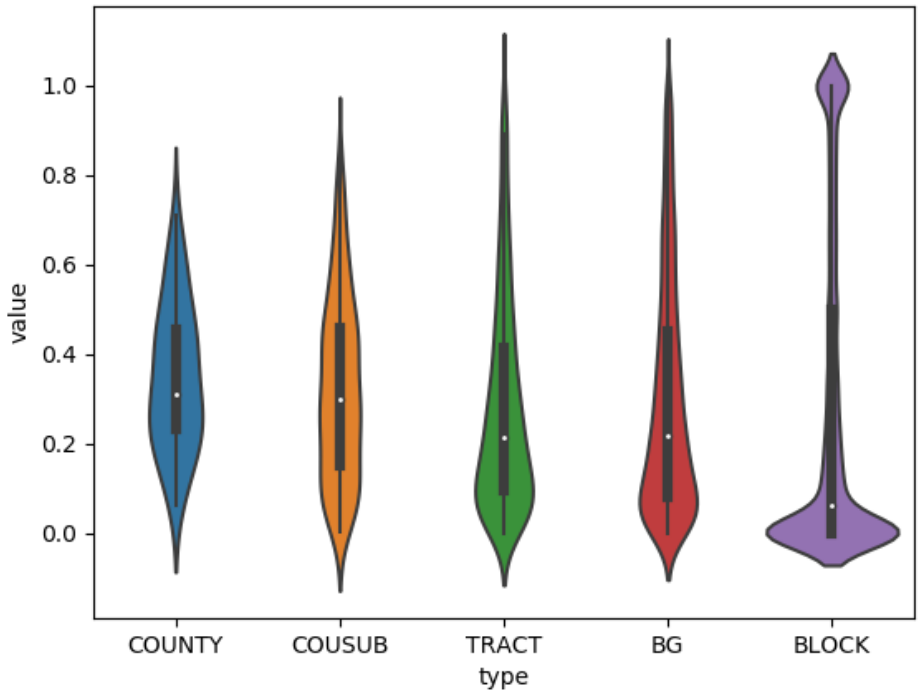
Homogeneity - BVAP - Louisiana



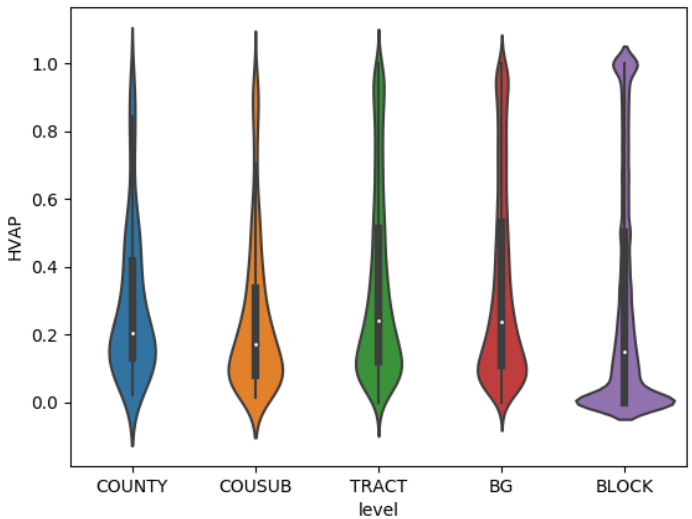
Homogeneity - BVAP - Mississippi

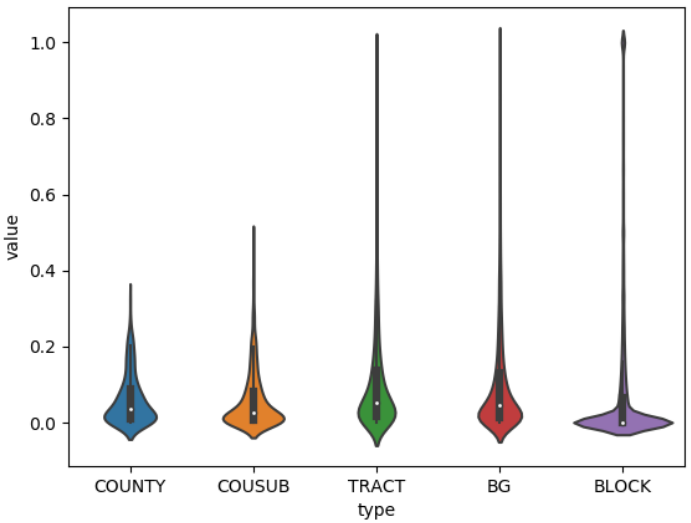


Homogeneity - BVAP - South Carolina

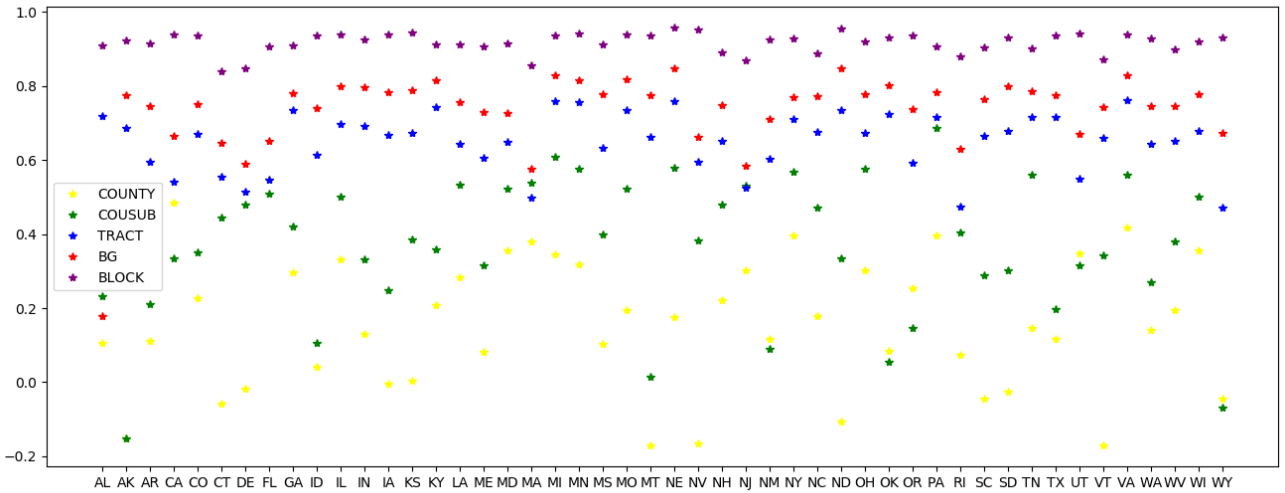


Homogeneity - HVAP - BVAP - Texas

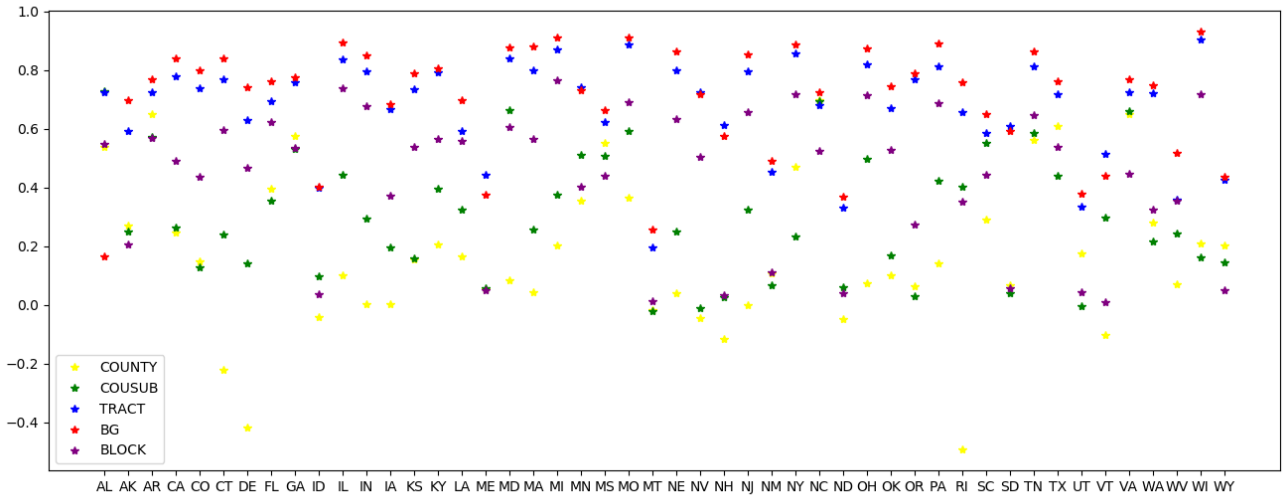




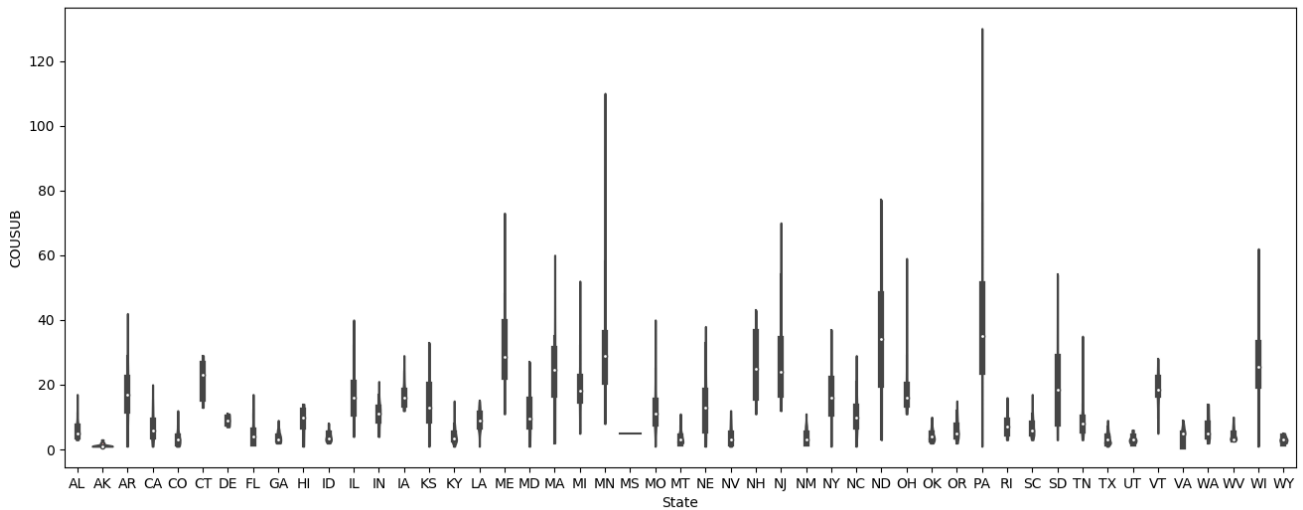
Assortativity (Urban/Rural)



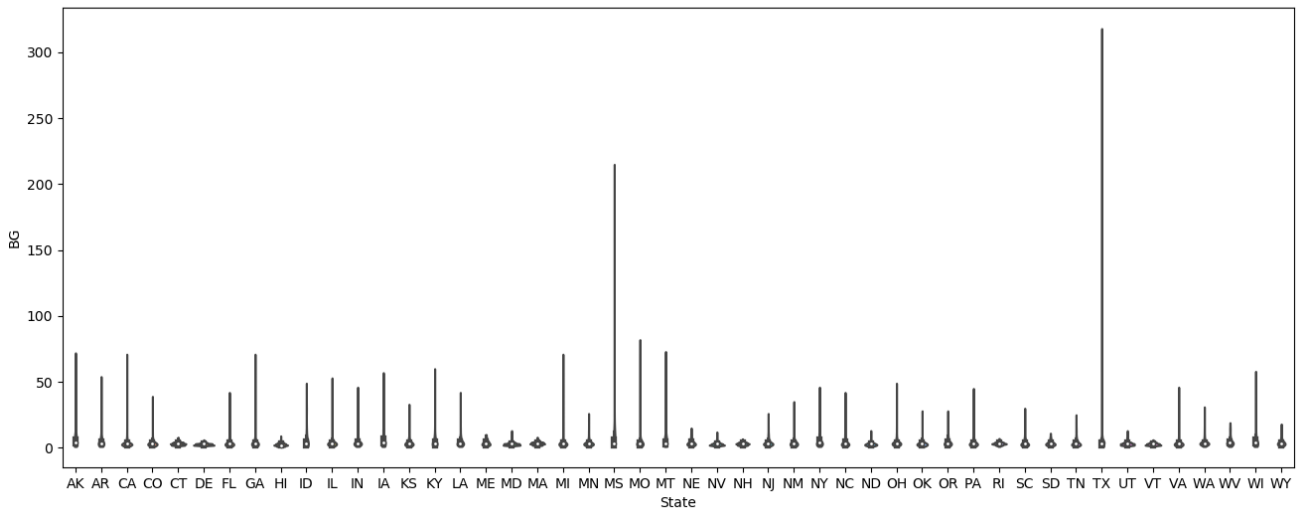
Assortativity (BVAP)



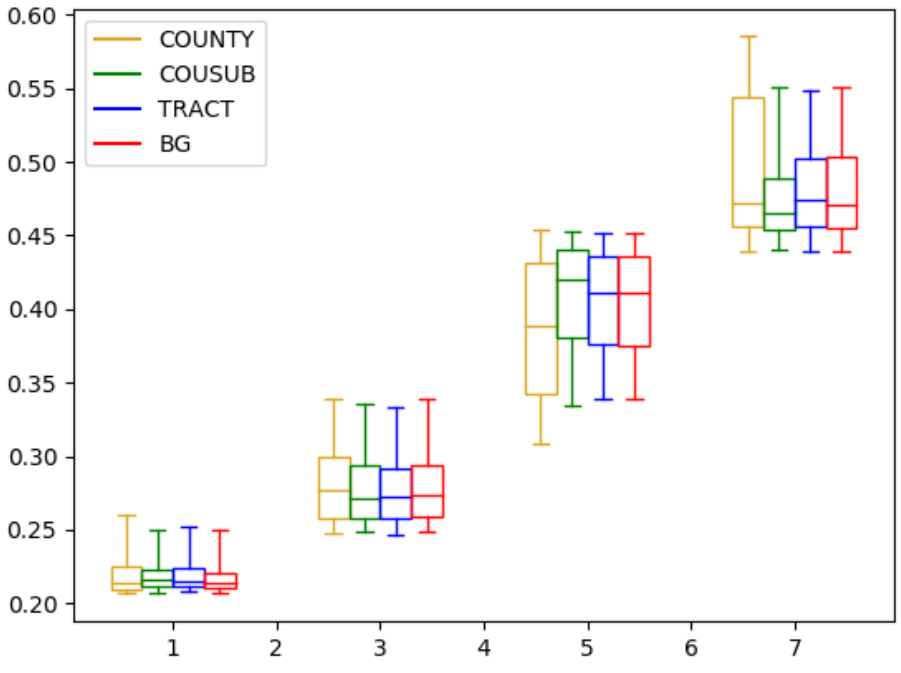
Nesting


COUSUB
0
20
40
60
80
100
120
AL AK AR CA CO CT DE FL GA HI ID IL IN IA KS KY LA ME MD MA MI MN MS MO MT NE NV NH NJ NM NY NC ND OH OK OR PA RI SC SD TN TX UT VT VA WA WV WI WY
State

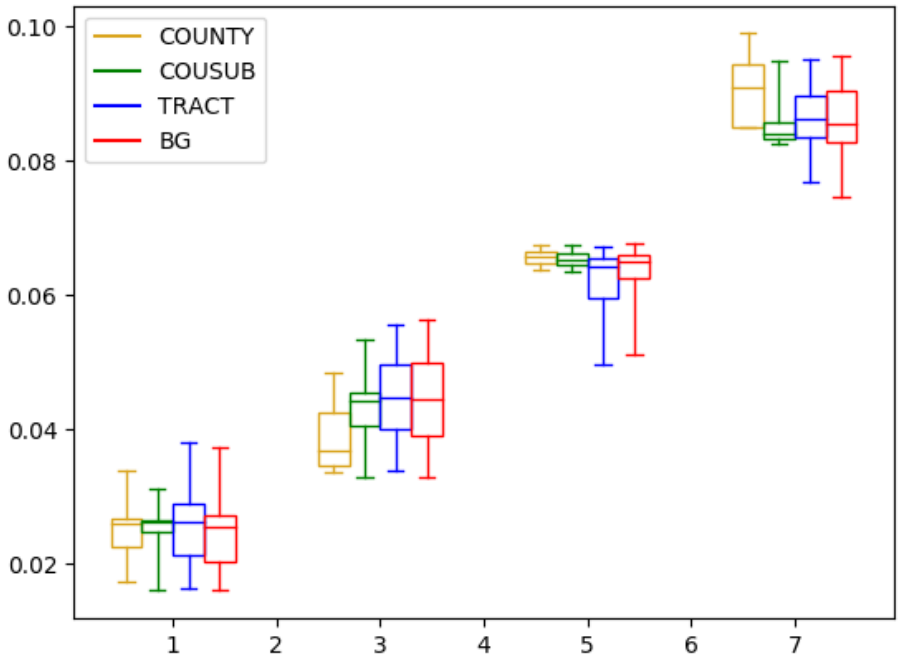
Nesting


BG
300
250
200
150
100
50
0
AK AR CA CO CT DE FL GA HI ID IL IN IA KS KY LA ME MD MA MI MN MS MO MT NE NV NH NJ NM NY NC ND OH OK OR PA RI SC SD TN TX UT VT VA WA WV WI WY
State

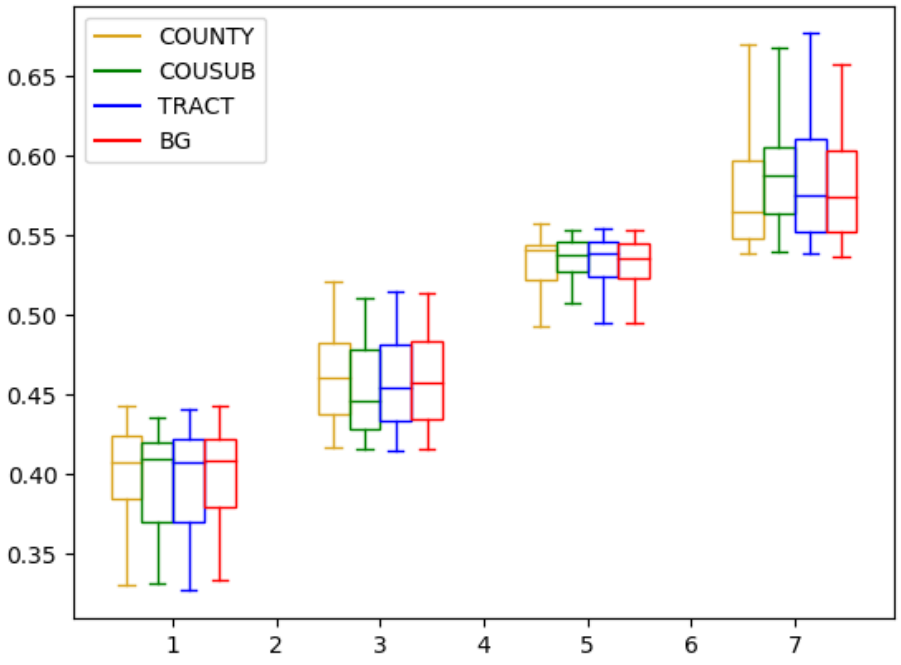
Ensemble Comparisons - BVAP - Mississippi



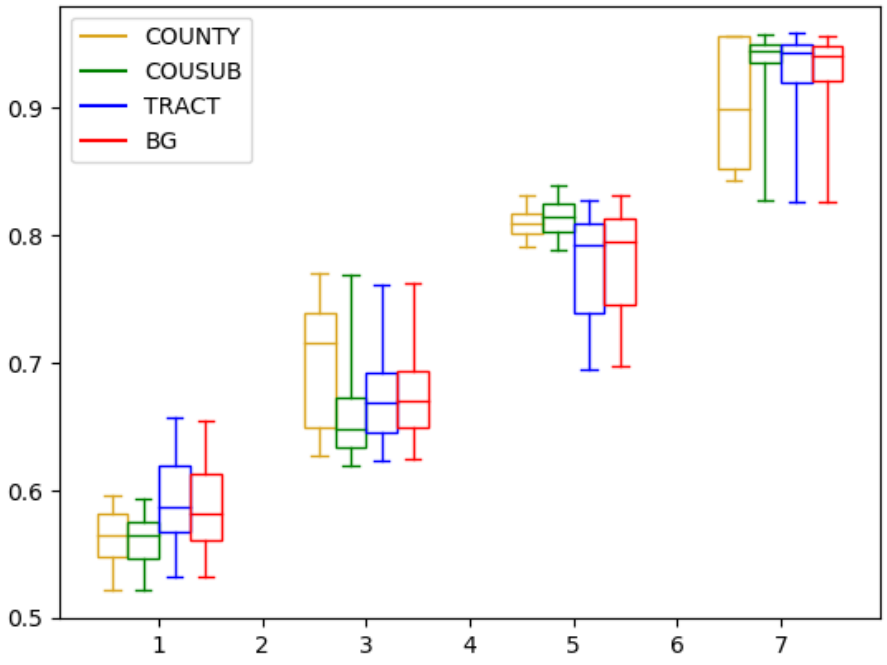
Ensemble Comparisons - BVAP - Kansas



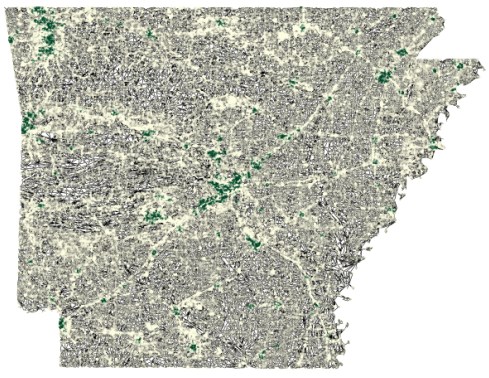
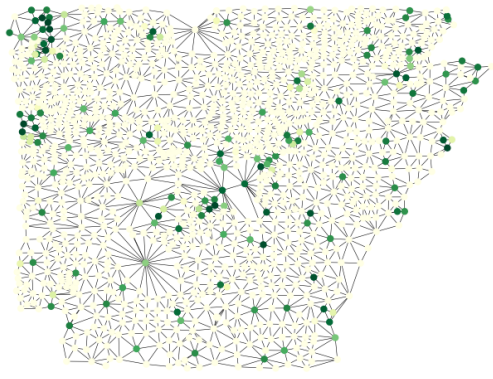
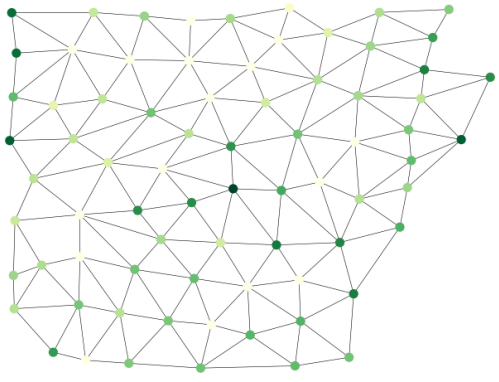
Ensemble Comparisons - Urban/Rural - Mississippi



Ensemble Comparisons - Urban/Rural - Kansas

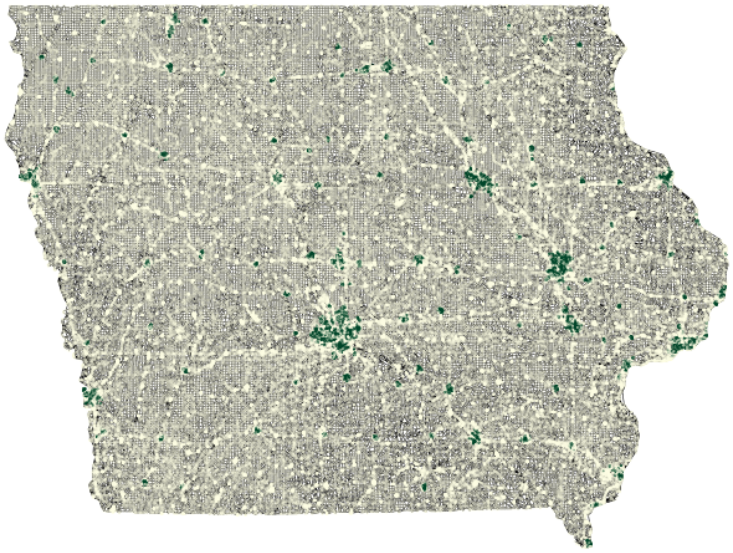


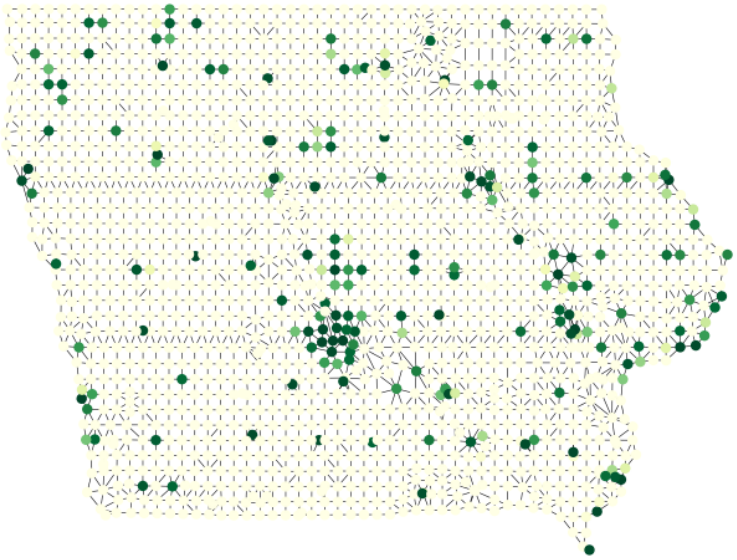
Census Rules

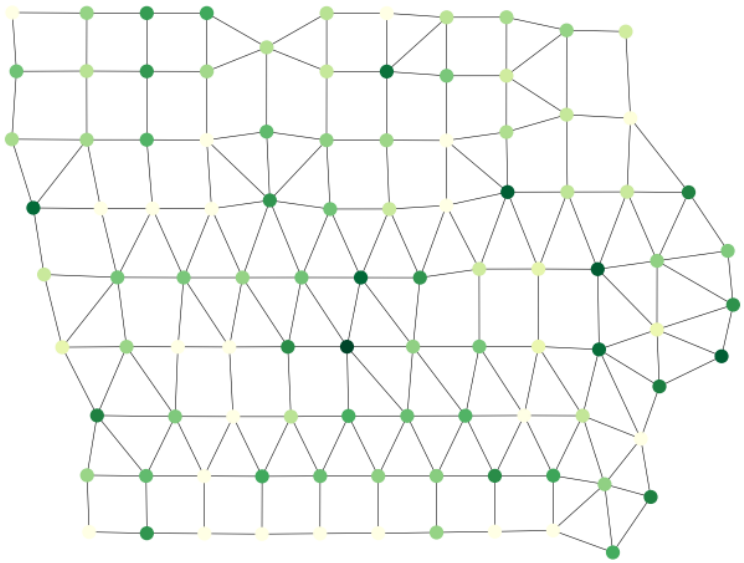
- Defined by population density and total population in contiguous collections of block groups
- Distinguishes between "hop" and "jump" closeness
- Full Guidelines: `https://www2.census.gov/geo/pdfs/reference/fedreg/ua_2k.pdf`

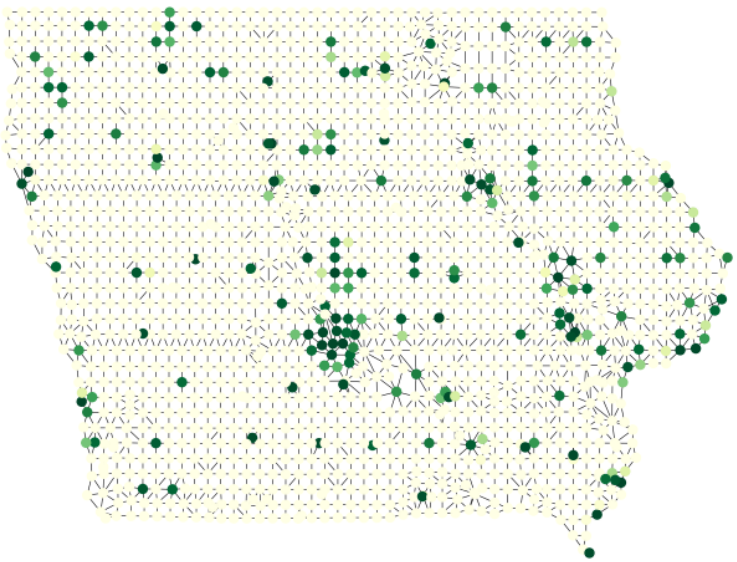
Examples



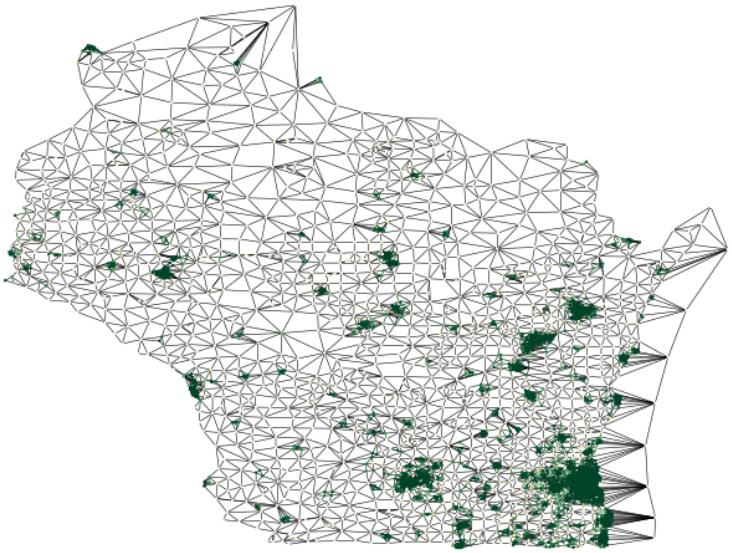


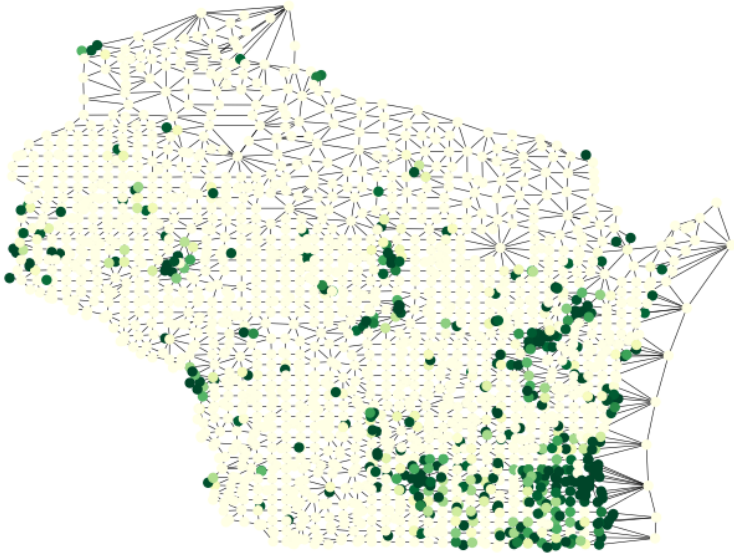
Examples



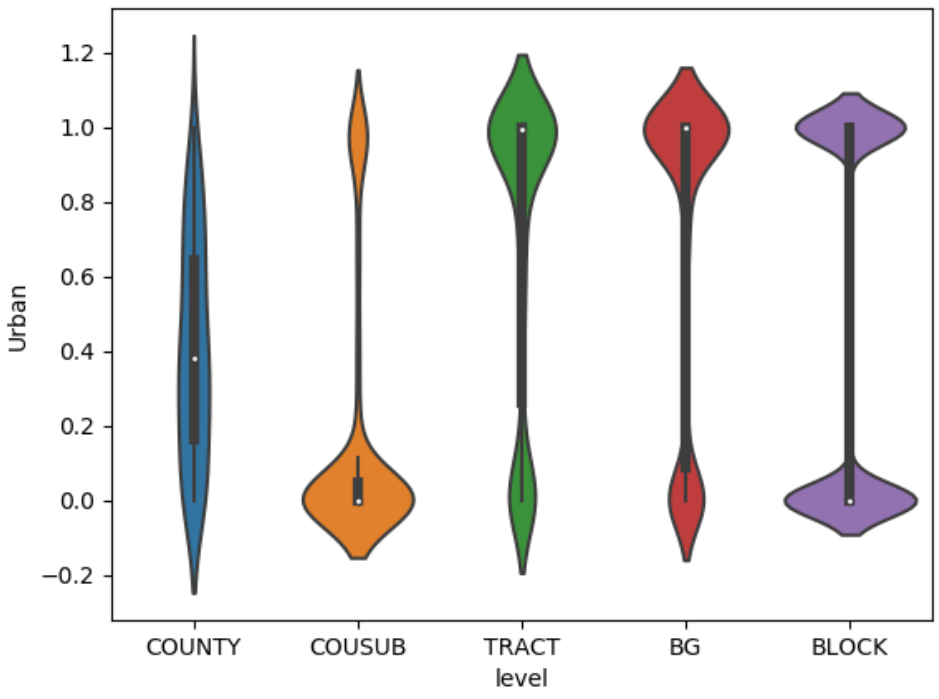


Examples





Homogeneity



Algorithmic Approaches

- Definitely Cheating:
  - Geographic/Geometric Data
  - Population or other demographic data

Algorithmic Approaches

- Definitely Cheating:
  - Geographic/Geometric Data
  - Population or other demographic data
- Partially Cheating:
  - Provide some set of initial labels (Belief Propagation)
  - Multiresolution (use nesting properties)

Algorithmic Approaches

- Definitely Cheating:
  - Geographic/Geometric Data
  - Population or other demographic data
- Partially Cheating:
  - Provide some set of initial labels (Belief Propagation)
  - Multiresolution (use nesting properties)
- Not cheating:
  - Extracting grid subgraphs
  - Centrality measures

Rules and Modeling

County Preservation

- Minimize to the extent possible
- Splitting bounds by county
- County Cluster

Rules and Modeling

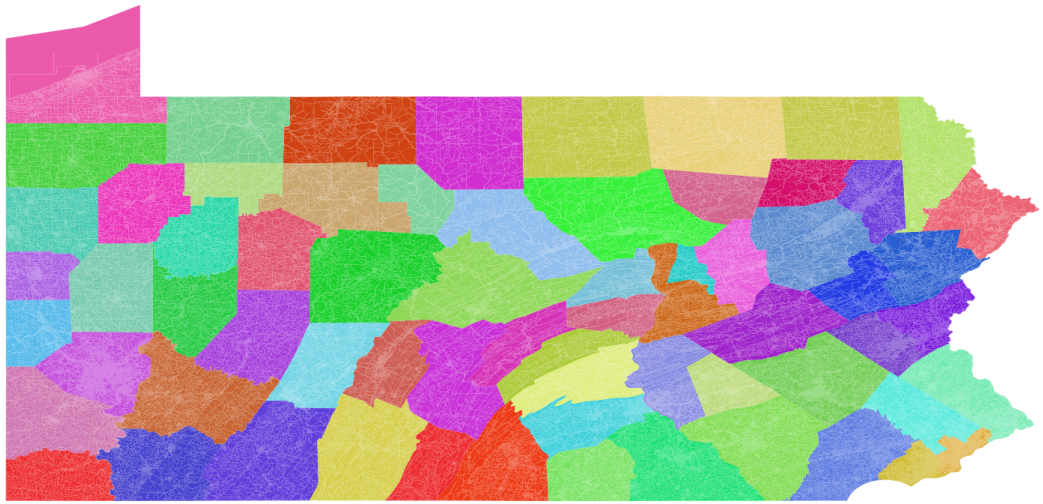
County Preservation

- Minimize to the extent possible
- Splitting bounds by county
- County Cluster

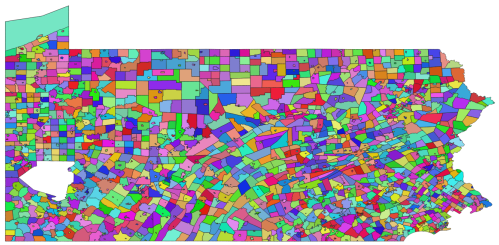
How to implement:

- Multi-resolution dual graphs
- Hierarchical Clustering
- Proposal Optimization
- Markov chains with marked edges

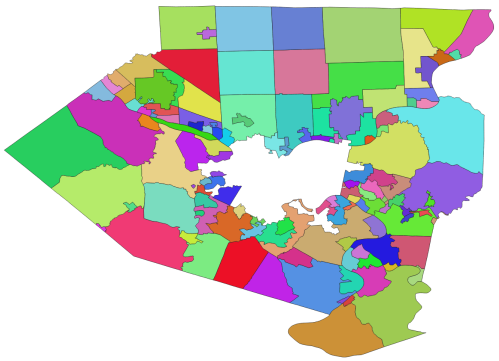
Counties



Municipalities

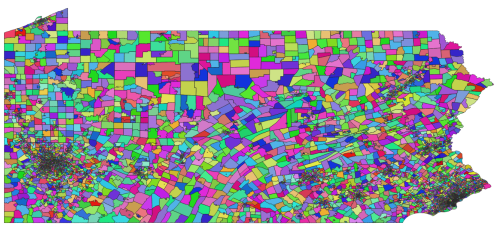


Pennsylvania

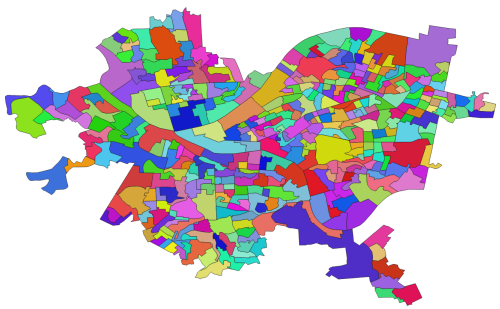


Allegheny

Precincts



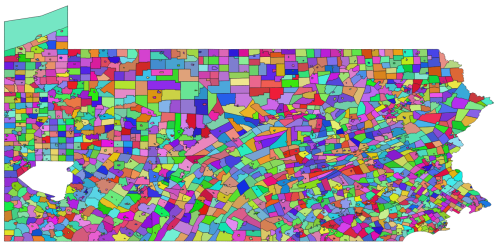
Pennsylvania



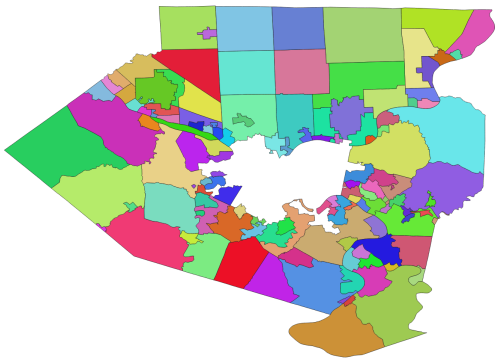
Pittsburgh

Philadelphia

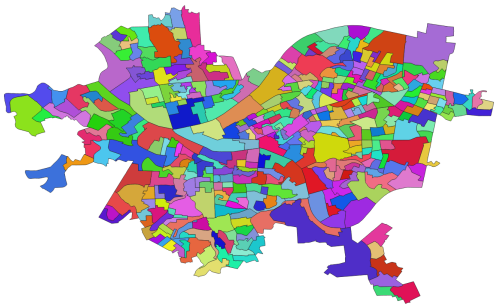
Putting Them Together



Pennsylvania



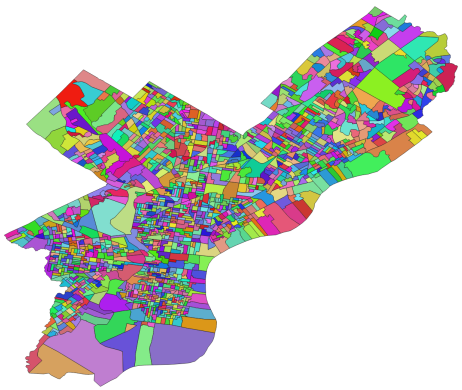
Allegheny



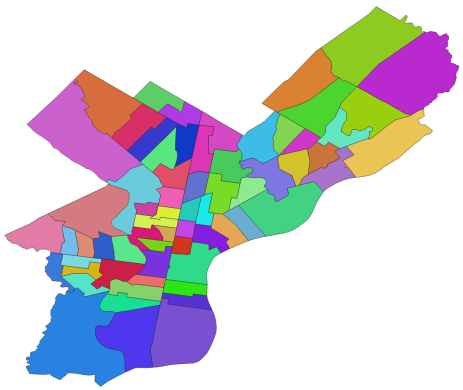
Pittsburgh

Putting Them Together

Blocks

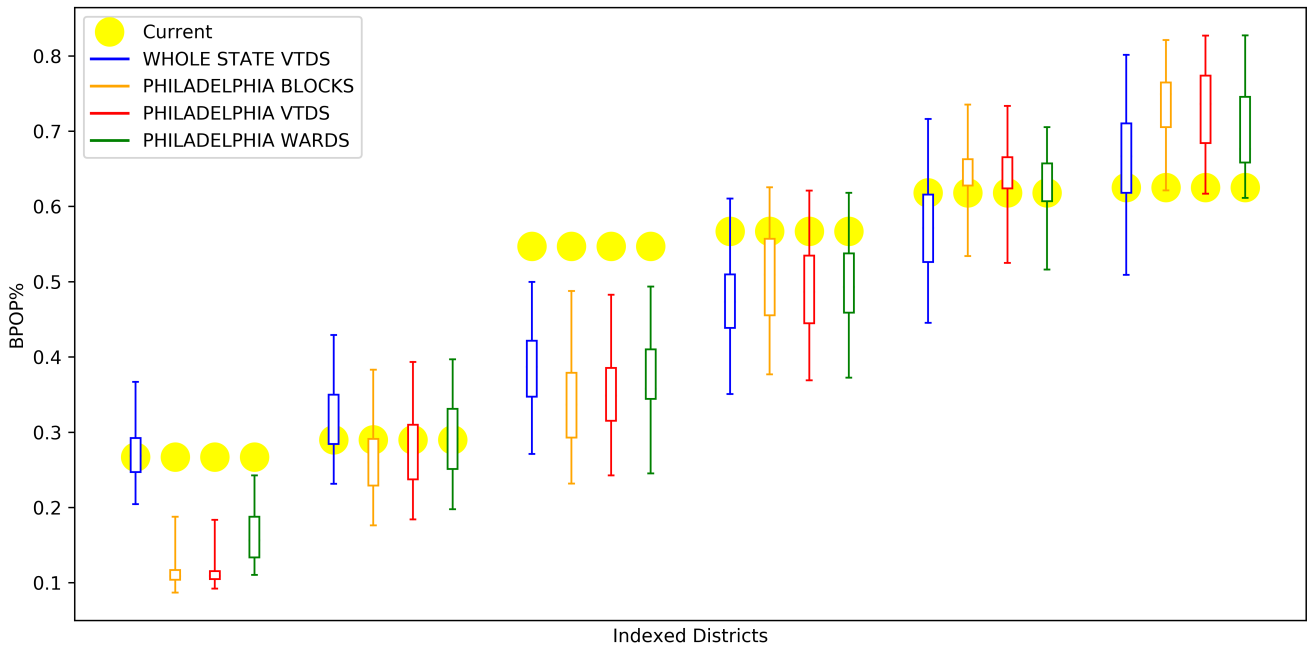


Precincts



Wards

Comparison of BPOP Districts



Building Seeds

- County Splits
- VRA compliance
- Partisan Metrics
- Population Balance

Building Seeds

- County Splits
- VRA compliance
- Partisan Metrics
- Population Balance

- "Coerce" with acceptance functions
  - Bound kth district changes
  - Ladder methods
- Local Hill Climbing
- Proposal Tuning

Weighted Trees

- At each step of the ReCom chain weight the edges of the dual graph by a random function of county similarity
- Draw a maximal spanning tree
- Cut like normal

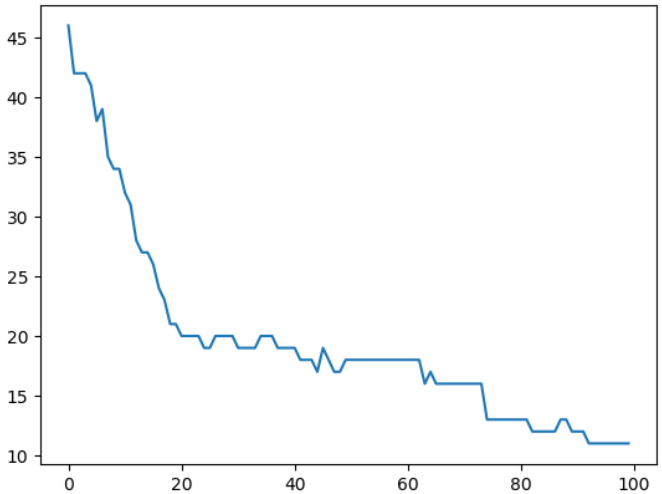
Weighted Trees

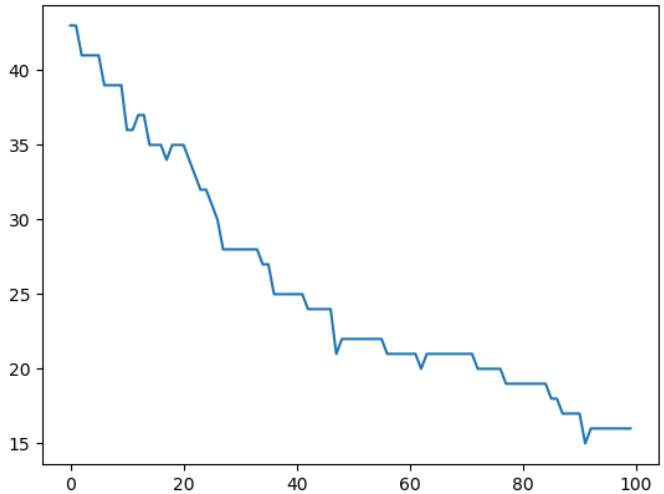
- At each step of the ReCom chain weight the edges of the dual graph by a random function of county similarity
- Draw a maximal spanning tree
- Cut like normal

Intuition

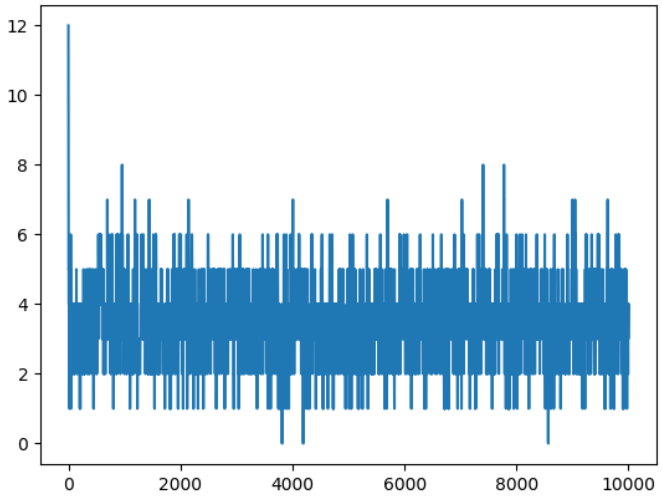
- In the limit, equivalent to drawing a spanning tree for each county and then a spanning tree on the county dual graph
- Imagine starting in one corner of the graph and following the tree through the counties
- The population constraint still requires some counties to be split

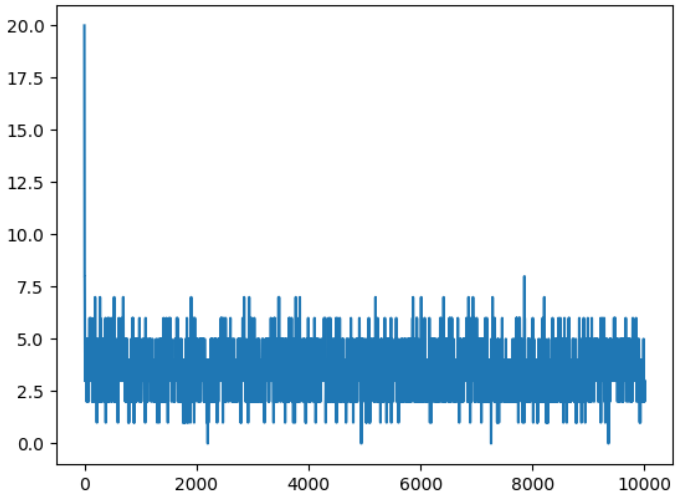
Virginia - Pennsylvania





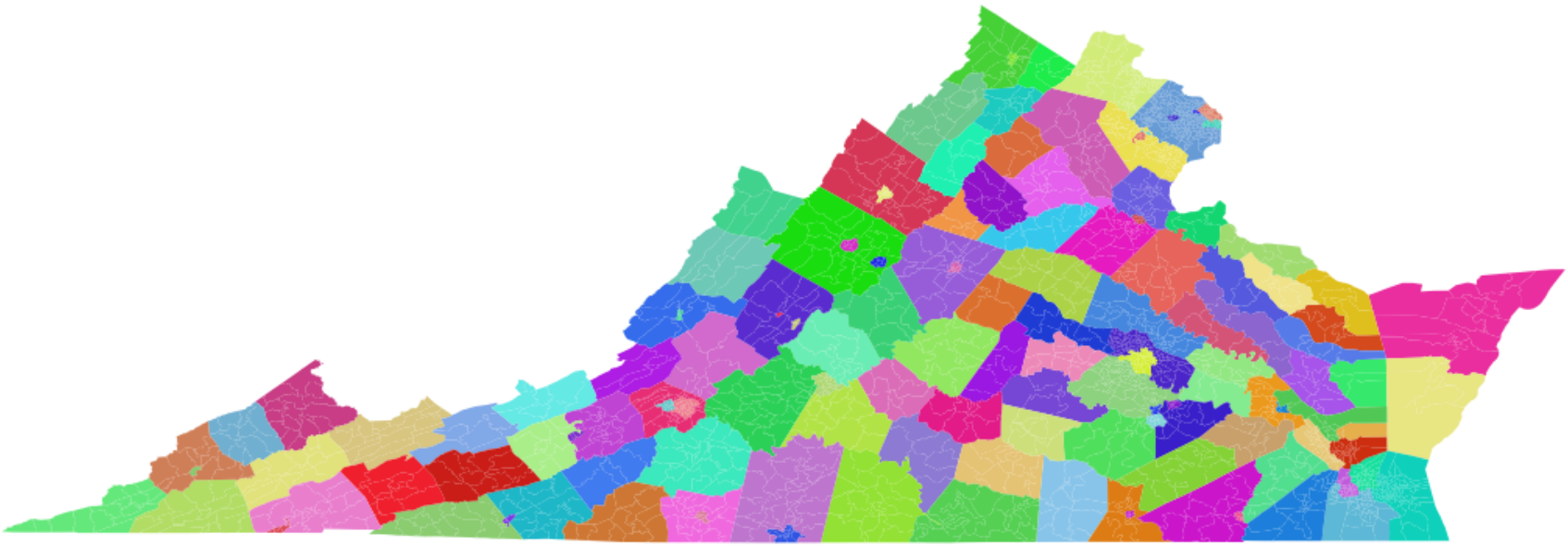
Kansas - Block Groups - County Subunits



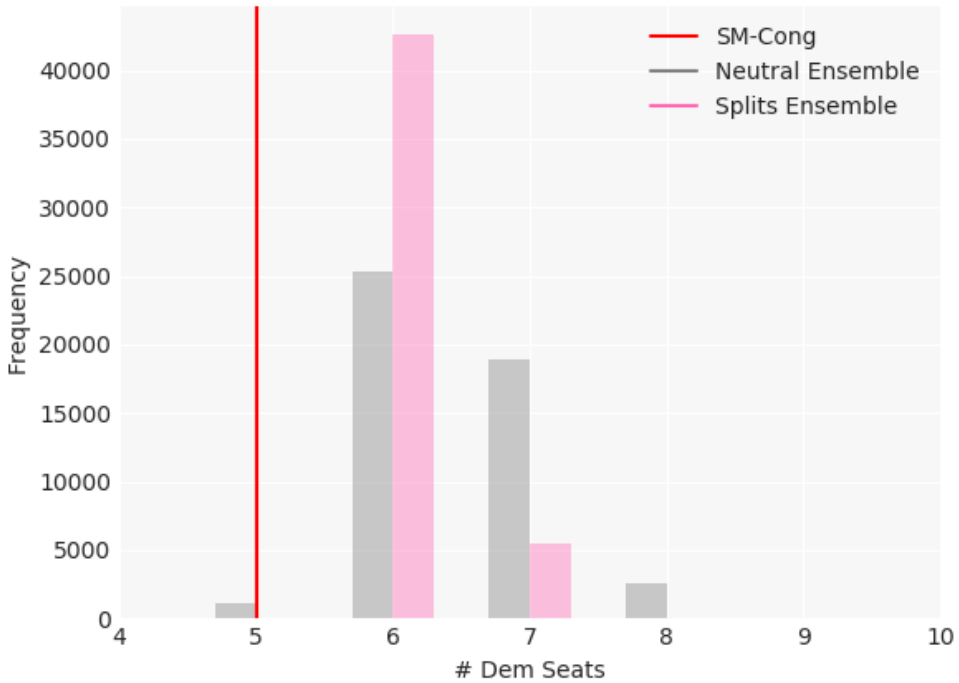


Animation

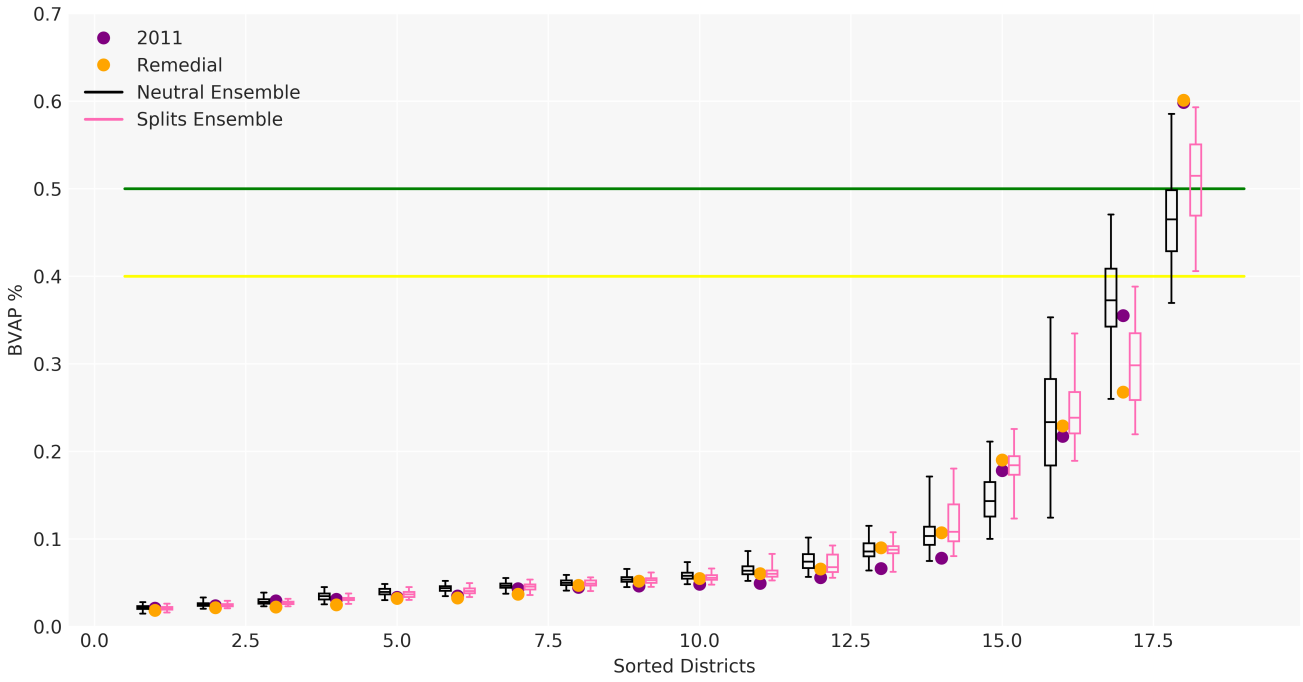
Virginia Localities



Partisan Impacts



Mixed Evidence



Questions?

Thanks![1]

[1] Computational Redistricting is not a solved problem! ☺